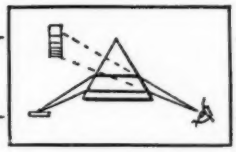
Journal of the OPTICAL SOCIETY *of* AMERICA

VOLUME 25
No. 9



SEPTEMBER
1935

Light Absorption and Distribution of Atmospheric Ozone . . . . . RUDOLF W. LADENBURG 259

Some Phases of the Mitogenetic Ray Phenomenon . . . . . ALEXANDER HOLLAENDER AND WALTER D. CLAUS 270

Experimental Determination of the Visual Thresholds at Low Values of Illumination . . J. PRESTON CONNER AND ROBERT E. GANOUNG 287

On the Fundamental Optical Invariant, the Optical Tetrality Principle, and on the New Development of Gaussian Optics Based on This Law . . . . . M. HERZBERGER 295

New Recording Spectrophotometer . . . . . ARTHUR C. HARDY 305

Some Stigmatic Grating Mountings . . . . . HAROLD T. SMYTH 312

Modification of the Fabry-Perot Interferometer . . . . . R. C. MACHLER AND RUSSELL A. FISHER 315

Published monthly
at Prince and Lemon Sts., Lancaster, Pa.
by the
AMERICAN INSTITUTE OF PHYSICS
Incorporated
11 East 38th Street, New York, N. Y.
FOR AND UNDER THE EDITORIAL SUPERVISION OF
THE OPTICAL SOCIETY OF AMERICA

THE OPTICAL SOCIETY OF AMERICA

OFFICERS OF THE SOCIETY

W. B. RAYTON, *President*
Bausch & Lomb Optical Company
Rochester, New York

ARTHUR C. HARDY, *Vice President*
Massachusetts Institute of Technology
Cambridge, Massachusetts

L. B. TUCKERMAN, *Secretary*
Bureau of Standards
Washington, D. C.

HENRY F. KURTZ, *Treasurer*
Bausch & Lomb Optical Company
Rochester, New York

The Board of Directors consists of the above officers; the Editor of the Journal; the Past President; and five elected members:

E. C. CRITTENDEN, *Past President*
Bureau of Standards
Washington, D. C.

F. K. RICHTMYER, *Editor*
Cornell University
Ithaca, New York

Elected Members

ARTHUR H. COMPTON
Ryerson Physical Laboratory
University of Chicago
Chicago, Illinois

HERBERT E. IVES
463 West Street
New York, New York

CARL W. KEUFFEL
Keuffel & Esser Company
Hoboken, New Jersey

GEORGE B. PEGRAM
Columbia University
New York, New York

R. C. GIBBS
Department of Physics
Cornell University
Ithaca, New York

THE OPTICAL SOCIETY OF AMERICA was organized in 1916 "*to increase and diffuse the knowledge of optics in all its branches, pure and applied, to promote the mutual interests of investigators of optical problems, of designers, manufacturers and users of optical instruments and apparatus of all kinds and to encourage cooperation among them.*" It aims to serve the interests of all who are engaged in any branch of optics from fundamental research to the manufacture of optical goods.

MEMBERSHIP

Regular Members: "Any person who has contributed materially to the advancement of optics is eligible for regular membership."

Associate Members: "Any person who is interested in optics is eligible for associate membership."

Corporate Members: "Any corporation desirous of encouraging the work of the Society is eligible for corporation membership."

The Society cordially invites to membership all who are interested in any branch of optics, either in research, in instruction, in optical or illuminating engineering, in the manufacture and distribution of optical goods of all kinds or in physiological and medical optics.

ANNUAL DUES

Regular members, $7.50; Associate members, $5.00; Corporate members, $50.00. These dues include subscription to the Society's official journal, the *Journal of the Optical Society of America.*

MEETINGS

The Society holds two meetings per year: An annual meeting in the autumn and a midwinter meeting jointly with the American Physical Society.

For further information concerning the Optical Society of America address the Secretary, Dr. L. B. Tuckerman, Bureau of Standards, Washington, D. C.

ROCHESTER, NEW YORK SECTION OF THE OPTICAL SOCIETY OF AMERICA

C. C. NITCHIE, *President*
H. E. SEEMANN, *Secretary*

L. A. DUBRIDGE, *Vice President*
H. G. KINGSLAKE, *Treasurer*

Councilors: E. C. FRITTS, R. KINGSLAKE, H. F. KURTZ, T. R. WILKINS, C. J. STAUD (*ex-officio*)

Journal of the
OPTICAL SOCIETY *of* AMERICA

VOLUME 25 SEPTEMBER, 1935 NUMBER 9

Light Absorption and Distribution of Atmospheric Ozone[1,2]

RUDOLF W. LADENBURG, *Palmer Physical Laboratory, Princeton University*

(Received May 25, 1935)

THE spectrum of the sun does not reach as far into the infrared or the ultraviolet as it should if it were a black body of 6000°. The short wave-lengths especially are cut off, the spectrum ends very abruptly below 3000A. Its intensity relative to 4000A should be 61 percent at 3000, 9 percent at 2000A, but there is nothing below 2900A. In fact, the end of the spectrum occurs at even longer wave-lengths in the morning and evening, i.e., when the sun is low. Cornu, the pioneer worker in the field, concluded as early as 1878 that absorption in the atmosphere of the earth is the reason for the short wave-length limit of the sun spectrum, the path of the sunlight in the atmosphere being obviously the longer the lower the sun stands. Shortly later Hartley discovered the strong absorption of ozonized oxygen in the region between about 3200 and 2300A, the "Hartley band," and suggested that ozone is the medium absorbing the short wave-length end of the sun spectrum. This end is, under favorable conditions, independent of the accessible height above sea level: Miethe and Lehmann (using in their different experiments the same spectrograph with crossed quartz prisms in order to avoid stray light of the immensely stronger longer wave-lengths) got practically the same short wave-length limit (2913A) in Berlin and in different places in Switzerland up to a height of 15,000 feet above sea level (compare Table I); and even at the twofold height of 30,000 feet Wiegand in a balloon ascent could only reach a few Angstrom units more. The record so far has been attained by F. W. P. Goetz in Arosa at the height of 6000 feet, who got a picture down to 2863A using a fast spectrograph and long exposure time and avoiding every trace of stray light. If it is ozone which hinders us from studying the short waves of the spectrum of the sun and the stars, this is certainly high in the stratosphere. The investigations of Fabry and Buisson in 1913 and the

TABLE I. *Shortest wave-length of sun spectrum.*

	Height above sea level	Date	Wave-length
Cornu:			
in Courtenay	200 m	June 24, 1878	2930A
Miethe and Lehmann:			
in Assuan	116 m	February 1908	2916
Berlin	50	Sept. 8, 1908	2913
Zermatt	1620	14, 1908	2914
Gornergrat	3136	24, 1908	2911
Monte Rosa	4560	27, 1908	2912
Wiegand:			
balloon ascent	9000 m	Sept. 9, 1912	2896
Fabry and Buisson:			
in Marseille	10 m	June 20, 1920	2885
F. W. P. Goetz:			
in Arosa	1900 m	July 9, 1930	2863

[1] Presented by invitation at the Symposium on Atmospheric Optics held under the joint auspices of the American Physical Society and the Optical Society of America, February 22, 1935.

[2] The literature up to 1930 and many details of this subject are to be found in the reports of C. Fabry and H. Buisson, *Mémorial des Sciences Physiques de l'Acad. des Sci. de Paris*, Vol. XI, 1930, and of F. W. Paul Goetz, *Ergebn. d. kosm. Physik*, Vol. I. Leipzig 1931.

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

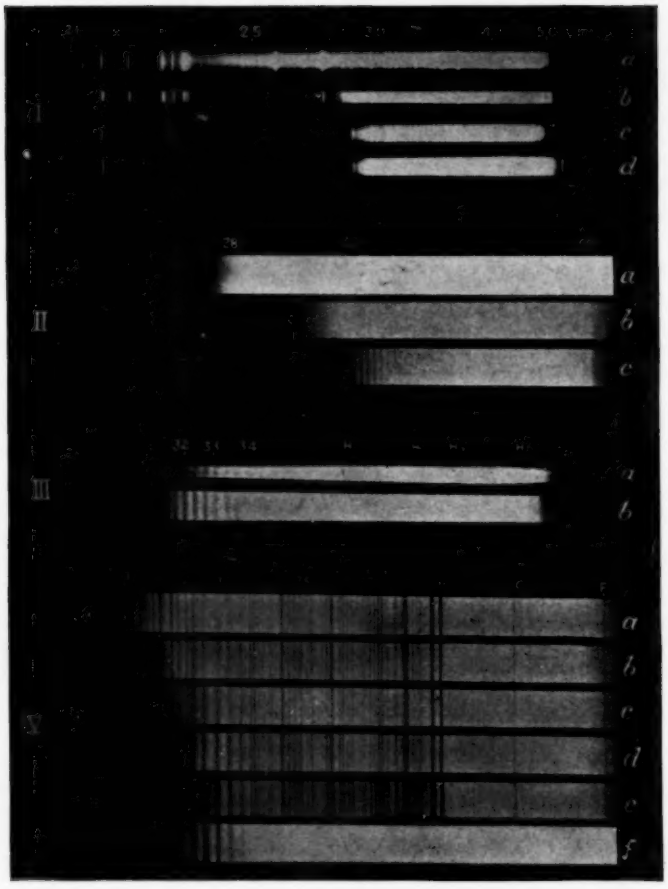


FIG. 1. Absorption spectrum of ozone in the ultraviolet (I and II) and sun spectrum (IV).

spectrograms obtained by A. Fowler and R. J. Strutt[2a] in the year 1917 prove without doubt that it *is* the ozone (compare Fig. 1).

The pictures I (Fig. 1) show the ultraviolet spectrum of a cadmium spark (a) without ozone and (b, c, d) absorbed by increasing amounts of ozone. Pictures II a, b, c illustrate the advance of the great Hartley band towards the red and show the appearance of a discontinuous absorption band spectrum above 3200A, discovered by E. Ladenburg and E. Lehmann in concentrated ozone; the pictures are taken with burning magnesium as light source traversing increasing thicknesses of dilute ozone (1 cm, 45 cm and 175 cm). The same absorption bands appear in picture III b and are also clearly visible in III a, the spectrum of the Sirius photographed in Edinburgh on February 8, 1915 at a zenith distance of the star of about 75°. They appear again in the spectra of the sun (IV a–e) in which the zenith distances changes from 45° to 89°; IV f gives again the ozone absorption bands as comparison.

[2a] A. Fowler and R. J. Strutt, Proc. Roy. Soc. **A93**, 577 (1917).

Without the presence of ozone in the stratosphere, life on earth probably would not exist in its present form. The ultraviolet radiation from the sun exerts strong physiological effects upon every organism, as for example formation of erythems, destruction of bacteria, production of vitamins. The very shortest wave-lengths of the sun have the greatest effect on the human skin. Small changes of the intensity of the sun in this region of the spectrum are biologically very efficient. These facts account for the wide interest

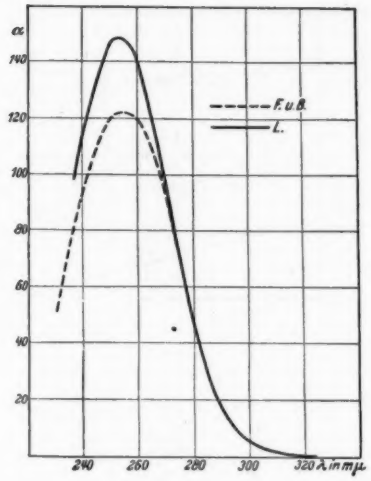


FIG. 2. Absorption coefficient α of ozone in the ultraviolet according to Ch. Fabry and H. Buisson (F. and B.), J. de physique (5) **3**, 196 (1913); and to A. Läuchli (L.) Zeits f. Physik **53**, 92 (1929)).

and the general importance of the study of the atmospheric ozone.

The first quantitative investigations have been carried out by Fabry and Buisson. They measured first (1913) the absorption coefficient of ozone for different wave-lengths in the Hartley band, which increases very rapidly between 3100 and 2900A, and later (1921) the intensity distribution of the sun spectrum in the same region at different zenith distances. By subtracting the well-known Rayleigh scattering they got an absorption curve of the sun spectrum being identical at low wave-lengths with that of an amount of 0.3 cm of ozone of NPT, that means if the whole amount of ozone in the atmosphere were compressed and reduced to 760 mm and 0°C it would give a layer of 3 mm.

Fig. 2 shows the change of the absorption coefficient of ozone in the Hartley band, when the lines of an Hg arc are used as light source, a being defined by the equation[3] that the transparency of a layer of l cm ozone at NPT

$$D = I/I_0 = 10^{-al}.$$

In such measurements ozone was always mixed with a much larger amount of oxygen. If one uses as light source the continuous spectrum of a hydrogen discharge at suitable conditions, one finds an ordinary absorption band

[3] In Fig. 2 the Greek letter α is used instead of a.

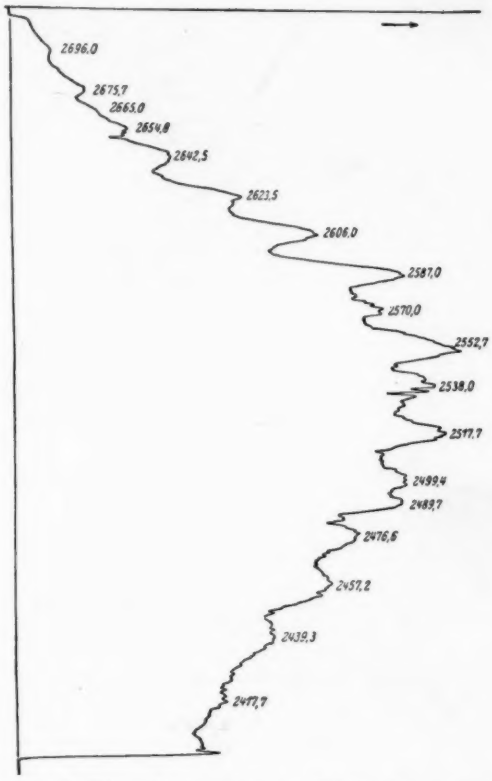


FIG. 3. Discontinuous structure of absorption spectrum of ozone in the ultraviolet (according to M. Lambrey and D. Chalonge, Gerlands Beitr. Geophysik. **24**, 45 (1929)).

spectrum with many maxima and minima (compare Fig. 3). It is obvious that the coefficient of absorption will depend to a large extent upon the total gas pressure and the temperature of the gas mixture and to some extent upon the breadth of the spectral lines used. It is therefore doubtful whether the coefficients as measured hitherto will allow a determination of the exact amount of the ozone under its conditions in the atmosphere. It seems that new measurements are necessary which are adapted in a better way to the atmospheric circumstances.

Fig. 2 gives besides the measurements of Fabry-Buisson (F. and B.) the more recent ones of Läuchli (1928), which are somewhat different near the maximum at 2550A. Here the absorption is as strong as that of metals: a layer of gaseous ozone of only 2/1000 cm thickness reduces the intensity to 1/2.[4] Table II gives the trans-

[4] A similar amount of absorption is effected by oxygen at 1450A, in the center of its absorption band in the Schumann region; the shape of this band resembles very

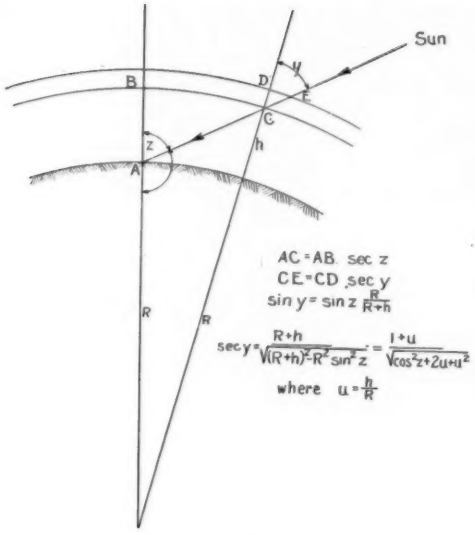


FIG. 4. Light path of the sun in the atmosphere.

TABLE II. *Absorption coefficient α and transparency D of 3 mm ozone at NPT for different wave-lengths.*

λ =	2300	2400	2500	2600	2700	2800	2900	3000	3100	3200	3300A
α =	50(?)	95	120	120	100	48	15.5	4.4	1.32	0.39	0.11
D =	10^{-15}	10^{-28}	10^{-36}	10^{-36}	10^{-30}	10^{-14}	2.10^{-5}	0.048	0.40	0.76	0.93

2×10^{-5}

parency D of a layer of 3 mm ozone at NPT for different wave-lengths. The most important feature for the study of the amount of ozone in the atmosphere is the rapid decrease of D from 0.4 at 3100 down to about 2.10^{-5} at 2900A.

Consider on the other hand the light absorption in the atmosphere. If k is its coefficient of total absorption plus scattering, I_0 the light intensity outside the atmosphere, I the intensity observed and t the total length of light path,

$$\log I = \log I_0 - k \cdot t. \qquad (1)$$

Fig. 4 shows how the light path of the sun increases with increasing zenith distance z, namely proportional to sec z if the atmospheric layer considered is near the earth surface, and proportional to sec y if this layer is in the height above the surface, where y is the sun's zenith distance seen from that height. Obviously

much that of ozone. (Cf. the values of its absorption coefficients, measured by R. Ladenburg and C. C. Van Voorhis, Phys. **Rev. 43**, 315 (1933).)

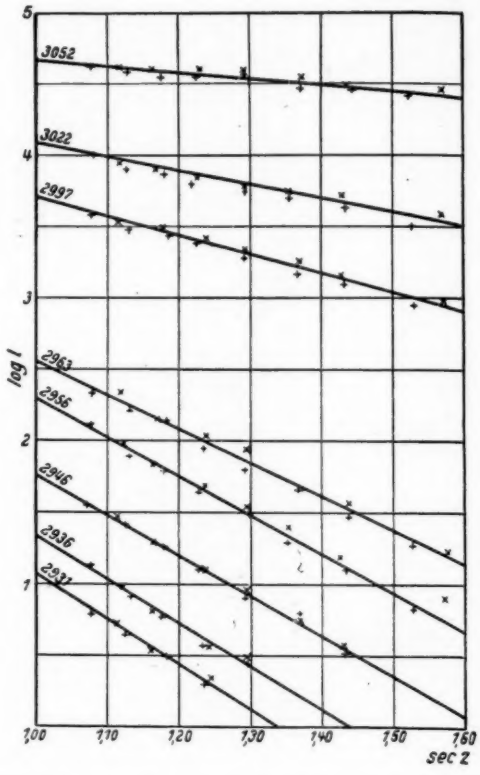


FIG. 5. Change of the intensity of the sun with zenith distance z and with wave-length according to Fabry and Buisson (J. de physique **62**, 197 (1921)).

$$\sin y = \sin z \cdot R/(R+h)$$

if R is the radius of the earth. For values of $h<60$ km and of $z<50°$ (sec $z<1.56$) is no appreciable difference between sec z and sec y; as long as $z<60°$ sec $y-\sin z<\sec z \cdot 1.05$.

If x is the total thickness of the ozone layer in the zenith, b the coefficient of Rayleigh scattering, H the height of the homogeneous atmosphere (about 8 km) and d the coefficient of scattering by dust or cloud particles etc.,

$$kt = ax \sec y + (b+d) \cdot H \sec z, \qquad (2)$$
$$b = (32\pi^3/3N\lambda^4)(\mu-1)^2 \cdot 0.434,$$

where N is Avogadro's number 2.70×10^{14}, μ the refractive index of the air, 0.434 the modulus of the logarithmus naturalis. d is very different for particles of different size. Even on very clear days and above the white snow of high plateaus (for

TABLE III. *Intensity of the sun on June 7, 1920 in Marseille.*

λ	k	Extraterrestrial	Zenith $z=0°$	$z=60°$
3143	0.84	155,000	22,400	3230
3104	0.99	147,000	15,100	1550
3052	1.40	256,000	10,200	410
3022	1.77	158,000	2,700	46
2997	2.22	218,000	1,320	8
2963	3.10	166,000	132	0.1
2956	3.33	162,000	76	0.04
2946	3.73	135,000	25	0.005
2936	4.12	141,000	11	10^{-3}
2931	4.36	126,000	6	10^{-4}
2922	4.82	145,000	2	10^{-5}
2917	5.08	105,000	0.9	10^{-5}
2912	5.39	74,000	0.3	10^{-6}
2906	5.78	25,000	0.04	10^{-7}
2898	6.36	45,000	0.02	10^{-8}

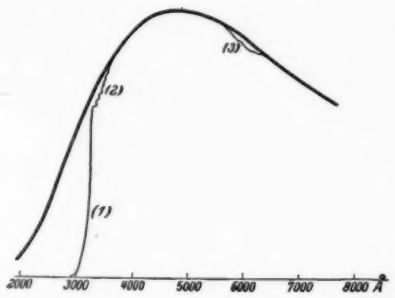


FIG. 6. Energy loss of sunlight between 2000 and 8000A produced by an amount of 3 mm ozone.

example in Arosa in Switzerland, 6000 feet high) there is always some scattering by particles, but experience shows that d changes only very slowly with λ and can be considered as constant in narrow spectral regions.

The procedure of Fabry-Buisson and of their followers is to determine I by photometric measurements for different z and for different λ in the critical region of rapid change of a (between about 3200 and 2920) and to plot the logarithm I/I' for two wave-lengths λ and λ′ as function of sec z. (As a matter of fact the ratio measured contains always a factor f depending on the spectrograph and the photographic plate, but this factor cancels in the calculation of k because

$$(\log (I/I'\cdot f)-\log (I_0/I_0'\cdot f) = \log I/I'-\log I_0/I_0').$$

As long as $z<50°$ one gets in this way straight lines (compare Fig. 5, where the ordinates give the log I/I' for λ = 3143A and for the wave-length written beside the lines). The slope of the lines gives $k-k'$. Assuming $d-d'=0$ and subtracting $(b-b')\cdot H$ one finds $(a-a')x$ and also x, provided that x does not change during the day. The fact that the values x for different wave-lengths turn out to be constant within the experimental errors of a few percent proves the usefulness of the method.

Very instructive is the reduction of the intensities to absolute values by comparison with a light source of known intensity distribution (determination of the factor f, mentioned above) and the calculation of the relative extraterrestrial intensities I_0 obtained from the intersection of the straight lines with the ordinate axis. Table III gives these values as calculated by Fabry-Buisson, which decrease in the region investigated only to 1/3, whereas the observed intensity of the sun at $z=60°$ decreases to less than 10^{-11}. The slow decrease of the extraterrestrial values corresponds in fact approximately to that of the radiation of a black body at 6000°. Fig. 6 shows graphically the energy loss of the sunlight (for $z=0$) in different spectral regions by an amount of 3 mm ozone, (1) refers to the Hartley band, (2) to the bands of Ladenburg-Lehmann and (3) to the weak Chappuis bands in the visible. In the ultraviolet the loss is about 4.9 percent of the total energy.

With measurements at low sun ($z>60°$) the values of log I/I' as a function of sec z do not lie on straight lines. By correcting for the effect of the Rayleigh scattering and by taking sec y instead of sec z, one should expect to get straight lines again corresponding to Eq. (2)—provided the hypothesis of a narrow ozone layer at a height h which determines the values of sec y has any meaning. According to the procedure of G. M. B. Dobson one draws a series of curves in which the abscissa sec z has been replaced by sec y corresponding to various assumed values of h (10 km, 20 km, 30 km, etc). On days with steady meteorological conditions one always finds that one of these curves is a straight line giving the value of h, which turns out to be in the inaccessible height of about 50 km. The slope of the lines gives the value of x in the same way as before.

Already the first measurements of Fabry-Buisson have shown appreciable changes of the total amount x from day to day; Goetz's longer series have demonstrated a definite seasonal influence and correlations to meteorological con-

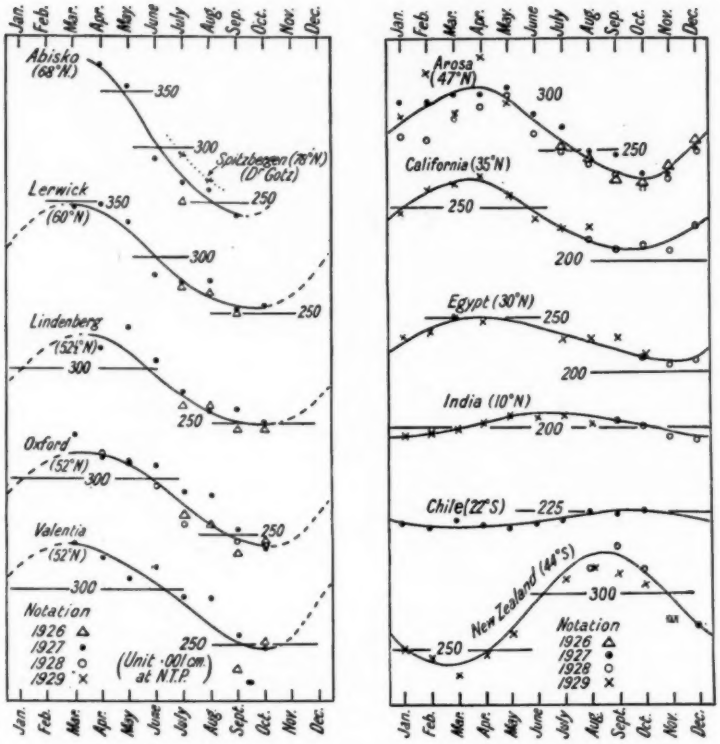


FIG. 7A and FIG. 7B. Variations of ozone during the year at different altitudes, according to a survey of G. M. B. Dobson (Proc. Roy. Soc. **A129**, 416 (1930)).

ditions. These important results attracted wide interest in different countries, and Dobson in Oxford constructed a large number of identical spectrographs designed especially for photographic photometric measurements of the sun spectrum near its ultraviolet end. He distributed these instruments in many observatories all over the world in order to get a continuous survey of the daily amount of ozone in different places of the earth, equally important for biology, geophysics and meteorology.

The most interesting results of this survey are contained in the following figures. Fig. 7A and 7B show the variation of ozone during the year for different latitudes. Everywhere the amount of ozone is high in spring and low in fall, the differences are greater the higher the latitude and almost disappear near the equator. Large changes have been observed, for example, in Lerwick at 60°N, the total amount being 0.417 in May and 0.214 in September. The high values after the long arctic night and their slow decrease during the long polar summer are somewhat surprising. It looks as if the sun had a more destructive than producing effect on the ozone. In fact, light absorbed in the great Hartley band destroys the ozone, while wave-lengths above and below 2000A dissociate O_2 into atoms forming O_3 in collision with oxygen molecules. Are there other agents responsible for the equilibrium between O, O_2 and O_3 and the atmospheric distribution of the ozone, as for example the corpuscles from the sun which dissociate the oxygen and produce the aurora? We will not discuss these questions as the report[46] of O. Wulf on photochemical reactions in the atmosphere treated this problem in some detail.

Another interesting correlation of the amount of ozone is that with the air pressure: low pressures ("cyclonic regions") are connected with

[46] Wulf, J. Opt. Soc. Am. **25**, 231 (1935).

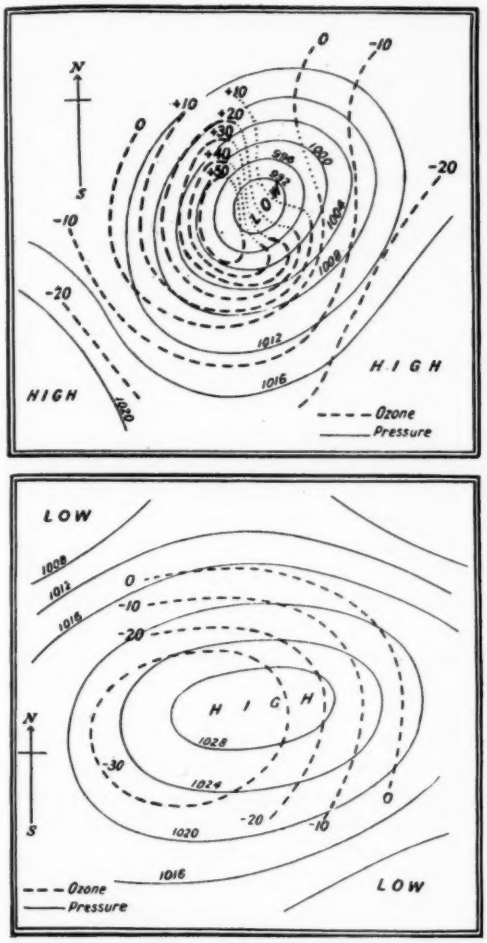


FIG. 8. Average variation of ozone with pressure (measured in millibar in cyclones and anticyclones, according to G. M. B. Dobson, D. N. Harrison and I. Lawrence (Proc. Roy. Soc. **A122**, 469 (1929)).

FIG. 9a. Station at 2500 m ("Hörnli").

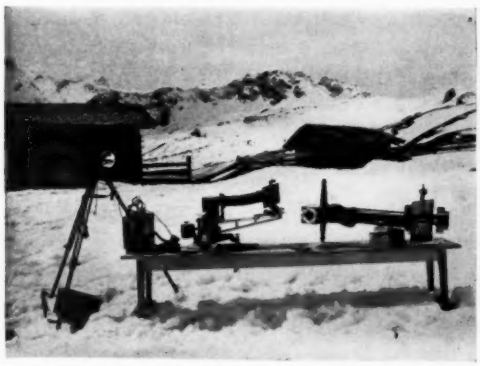


FIG. 9b. Station at 2000 m ("Tschuggen").

relatively large values of ozone, and *vice versa*, compare Fig. 8, where the pressure is given in millibar. The best correlation is established by the pressure at the top of the troposphere. This is difficult to understand if the ozone were concentrated in a narrow layer high in the stratosphere, a hypothesis doubted anyhow repeatedly on the ground of different theoretical considerations. Nevertheless, the estimations of its average height based on the method given above have always yielded values between 40 and 60 km, at least up to the year 1930. Then two investigations have changed completely our ideas about the distribution of atmospheric ozone: measurements of the amount of ozone near ground level and the determination of the height of ozone in Spitzbergen at a latitude of 80°N.

The observation of the mercury line 2537 over a distance of 6.5 km by R. J. Strutt in the year 1918 had shown that the amount of ozone near the surface of the earth certainly is very much smaller than would correspond with a uniform

distribution of ozone through the atmosphere; for 0.3 cm ozone distributed uniformly over 8 kilometers of the homogeneous atmosphere would give 0.26 cm in 6.5 km and would reduce the line 2537 to $1/10^{30}$ of its intensity. Since that time it had been generally assumed that practically no ozone was at ground level, as could be understood on account of the destructive action of organic substances on ozone. Then the real amount of ozone at the surface of the earth was measured photographically in the year 1930/1 by P. Goetz and R. Ladenburg near Arosa and by Buisson Jausseran and Rouard near Marseille. Both investigations yielded about 0.002 cm per km air. This amount is in approximate agreement with earlier chemical investigations which did not find due recognition on account of other chemical tests giving ten times more.

In such experiments spectra of a mercury arc are taken simultaneously with two spectrographs at different distances, in Arosa for example, from a distance of 1.2 and of 3.6 km; Fig. 9a and 9b show the two stations at 2500 m and 2000 m height above sea level and the spectrographs used. The amount of ozone in the air between the two stations can easily be determined from the density of the spectral lines, comparing the intensity ratio of the different wave-lengths with one and the same longer one and correcting in the usual way for the scattering of light. The measurements down to about 2600A yield a constant value of the amount of ozone. Below this wave-length another absorption sets in. This is mainly due to the band spectrum observed by Herzberg[5] in a layer of 25 m of pure oxygen. It has been confirmed by spectrograms taken with the continuous spectrum of a hydrogen arc as background by Goetz and Maier-Leibnitz, Chalonge and Bassy[6] who also compared the ozone content in Lauterbrunnen near Interlaken and on the Jungfraujoch (800 m and 3500 m above sea level).

These results proving the existence of an appreciable amount of ozone near the surface of the earth certainly do not correspond with a narrow ozone layer at a height of 50 km.

More recent measurements in Spitzbergen of the atmospheric ozone using the skylight of the zenith instead of the sunlight have changed the whole aspect. Such experiments have the advantage that they can be carried through at very low sun and even before sunrise and after sunset. They are not disturbed to any such extent as those using sunlight by stray light of longer wave-length, because the scattered light from the sky is relatively much richer in short wave-lengths. Using this method in Spitzbergen P. Goetz discovered the "Umkehreffect", which allows even the determination of the vertical distribution of ozone, at least to some extent.

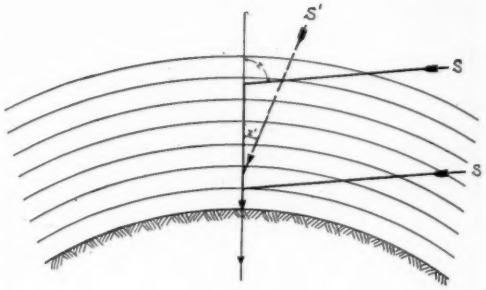


FIG. 10.

The amount of light reaching the measuring instrument which is scattered at any particular height depends on two things: the amount of air at that height available to scatter the light, and the absorption of light before and after scattering. The higher the scattering occurs the less light is scattered on account of the decreasing density of the air, but the less is also the absorption since the oblique path through the atmosphere is replaced more and more by the shorter vertical path (compare Fig. 10).

It can be shown that for each position of the sun and for each wave-length exists a limited region where most of the light is scattered. This effective height of scattering will be greater for smaller wave-lengths which are absorbed more rapidly, and will also be greater when the sun is low than when it is high. The process is somewhat involved, and it is not easy to oversee it quantitatively. But so much is clear: With increasing zenith distance the intensity scattered decreases because the average height increases and the path of the sunlight increases—therefore the intensity of every wave-length decreases and

[5] Herzberg, Naturwiss. 20, 577 (1932).

[6] Goetz and Maier-Leibnitz, Zeits f. Geophysik 2, 253 (1933); Chalonge and Bassy, J. de phys. et rad. (7) 5, 309 (1934).

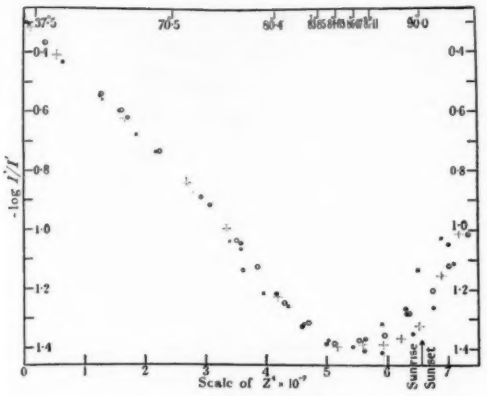


FIG. 11. Observed values of $-\log I/I'$ for zenith blue sky on 5 days on which the total ozone content was nearly identical. o, February 14, 1933, a.m. ($x=253$); ×, February 17, 1933, a.m. ($x=260$); ●, May 22, 1933, a.m. ($x=260$); +, June 3, 1933, a.m. ($x=263$); ○, June 3, 1933, p.m. ($x=256$). The "Umkehreffect."

more so the smaller the wave-lengths. If I corresponds to the shorter wave-length (3110A) suffering higher absorption ($a=$ *circa* 1.0) and I' to the longer one (3290A) suffering much less absorption ($a=$ *circa* 0.12), then the ratio I/I' decreases with increasing zenith distance z of the sun until the sun is rather low. Then, let us say, for $z=85°$ nearly all light of $\lambda=3110$ will be scattered above the ozone and a further increase of z does not decrease I much, but $\lambda'=3290$ is still scattered *in* the ozone region and will decrease more rapidly when the sun sets. Therefore I/I' raises again. Fig. 11, taken from a recent paper of Goetz, Meetham and Dobson,[7] shows this "Umkehr-effect" very clearly. Ordinates are the values $-\log I/I'$ measured with a new photoelectric instrument, especially constructed for this purpose by Dobson.[8] It allows much more accurate measurements than is possible by photographic photometry. As abscissae one uses not z values, but the fourth powers of z in order to open out the scale near sunrise and sunset. By careful mathematical analysis of such curves one can get some idea about the vertical distribution of ozone.

So far, no general analytical solution of the problem has been found. The simplest method applied is this: the atmosphere is divided into

[7] Goetz, Meetham and Dobson, Proc. Roy. Soc. **A145**, 416 (1934).

[8] Dobson, Proc. Phys. Soc. London **43**, 324 (1931).

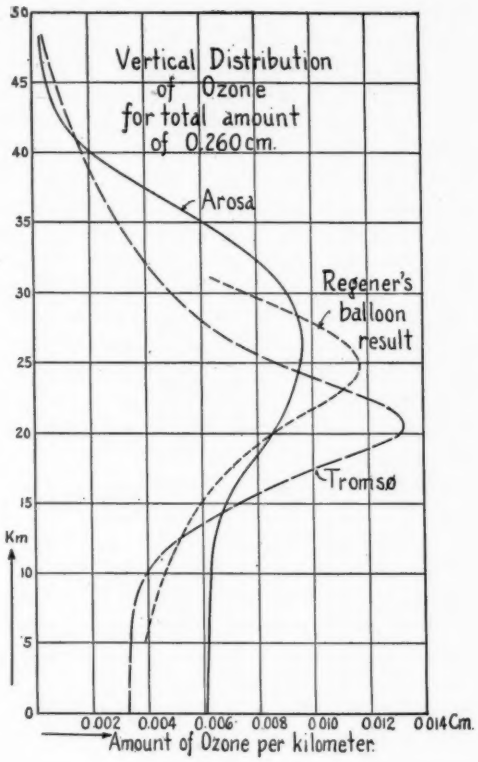


FIG. 12.

different layers, each containing a constant amount x_i of ozone; for example there are 5 layers assumed above Arosa: 2–5 km, 5–20, 20–35, 35–50 and above 50 km. In the last one x_5 is assumed to be zero. The amount x_1 in the first one is known from the measurements near ground level to be approximately one percent of the total amount of ozone. So there remain only two unknown values x_3 and x_4 whereas x_2 is taken as difference between x and the sum $x_1+x_3+x_4$. The result of such analysis is confirmed by more general calculations, assuming arbitrary amounts in 8 different layers and varying the assumed values till the theoretical "Umkehrcurve" shows a reasonably good agreement with the observed one. In general more than one such distribution can be found within the limits of error, but the "average height" of ozone is found to be almost identical for all such distributions.

Fig. 12 contains some of the results obtained so far, giving the amount of ozone in cm per km

vertical height in different heights above sea level. The curve "Arosa" is one of 6, published last fall by Goetz, Meetham and Dobson for different total amounts of ozone ($x=0.20$, $0.22\cdots0.30$ cm). The "average height" decreases in these cases from 23 down to 21 km. It is remarkable that similar values (about 27 km above sea level) have been obtained by simultaneous measurements with the new photoelectric instrument for the height of an assumed "narrow layer" according to the older method using sun light. The discrepancy with the earlier photographic photometric measurements ($h=$*circa* 50 km) is attributed by the authors mainly to the much greater accuracy of the new instrument. But there is no doubt that the "Umkehreffect" gives much more reliable results about the height of ozone than the older method using sun light: small diurnal variations of the amount of ozone affect the height deduced by this method very greatly.

Our Fig. 12 gives further a curve "Tromsø," which was kindly transmitted to me by Professor Dobson before their publication, based on more recent measurements carried out in Norway. It is also one of several curves showing similar differences as the "Arosa" curves for different total amounts of ozone.

The third curve, "Regener's balloon result," is taken from the first direct measurements of the amount of ozone in different heights up to 30 km obtained with sounding balloons by Erich and Victor H. Regener and published last October.[9] The agreement is very satisfactory, considering the totally different indirect method which led to the two other curves.

Regener's measurements have been carried out by means of a simple quartz spectrograph mounted in a light wooden frame shielded well against temperature changes and carried up by rubber balloons. The spectrograph was not focused upon the sun, but vertically down upon a disk coated with magnesium oxide which gives diffuse reflection of the sun and the sky light in all positions of the frame. An automatic device moves the photographic plate stepwise, and the plate is illuminated for some minutes between two steps. Pressure and temperature are simultaneously recorded on the same plate. Fig. 13 is an enlargement of the spectrograms showing the increase of the intensity of shorter wave-lengths with height. One measures the

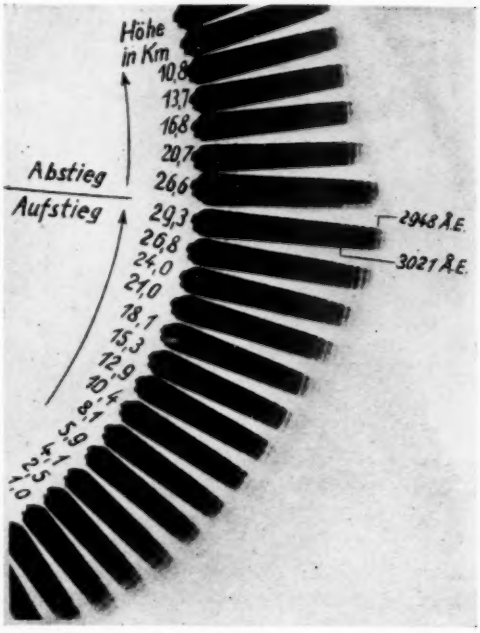


FIG. 13. Spectrogram of the Regeners taken in a sounding balloon.

density of the spectrogram and compares two wave-lengths from different heights giving the same density at the same exposure time. In this way one finds the amount of ozone above the instrument, starting at ground level and knowing the total amount of ozone. At the greatest height attained of about 31 km (at an air pressure of 9 mm Hg), the shortest wave-length measured with the relatively slow spectrograph and an exposure time of only 10 minutes was still 2875A, although 75 percent of the total amount of ozone was below the spectrograph (only about 0.06 cm ozone above); the maximum concentration of about 0.01 cm ozone per km was obtained at 24 km (compare Fig. 12).

The result obtained is very remarkable in different respects. It is definitely shown that the idea of a narrow ozone layer in 50 km height is wrong. The main part of ozone is not in inaccessible heights, but is distributed between 15 and 35 km with a distinct maximum at about 25 km, changing somewhat with the total amount of ozone. It is an important problem for the future to study this distribution at different seasons and on different places of the earth. The best method is certainly the direct measurement, using a spectrograph carried up by balloons following the pioneer work of the Regeners. Where this is not possible to do, the photoelectric instrument of

[9] Erich and Victor H. Regener, Physik. Zeits. **35**, 788 (1935).

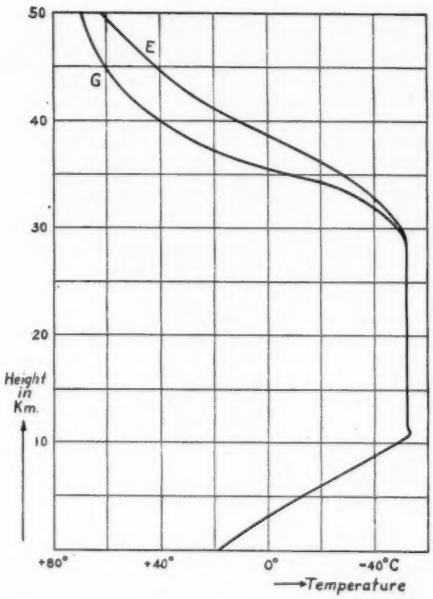


FIG. 14. Temperature of the stratosphere calculated from acoustical measurements. E = England (acc. to Whipple). G = Germany (acc. to Duckert). Taken from Bartels.[10]

Dobson has to be used, and Goetz's method of the "Umkehreffect" of the intensity of the sky light has to be applied. A more general mathematical solution of analysis should be worked out. The results of such a survey of the amount and distribution of ozone is of great importance on account of its correlation to meteorological, geophysical and biological data.

I may mention here the temperature distribution in the atmosphere. Formerly, the increase of the temperature in the region of 40–60 km height, deduced from the anomalous propagation of sound and the "zone of silence" surrounding large explosions,[10] was understood to be due to the absorption of sun energy by the ozone assumed to be concentrated in the same height. The new results show a nearly perfect correlation between the relatively high ozone concentration and the constant low temperature of −50°C in the stratosphere between 10 and 30 km (compare Fig. 14 taken from Bartels' report[10]). An understanding of this correlation will be possible probably in connection with the investigation of the optical properties and the equilibrium of O, O_2 and O_3, which should give simultaneously an explanation for the vertical distribution of ozone (compare the report of O. Wulf[46]).

The last and perhaps the most important problem involved in this matter is the question how it may be possible to get a knowledge of the ultraviolet spectrum of the sun and the stars below 2900, withheld from us so far by the ozone of the atmosphere. The remaining amount of 0.06 cm above 30 km still prevents the transmission of light between 2700 and 2400A—but it may be that the low values of ozone in autumn or near the equator will give surprising results. Some hope remains for the study of wave-lengths below 2400, where O_2 is mainly the absorber. One remarkable step forward has been recently made by E. Meyer, M. Schein and B. Stoll in Switzerland: Using a light counter[11] sensitive only to very short wave-lengths they were able to get a distinct small effect between 2200 and 2000 in Arosa and an appreciable increase of this effect on the Jungfraujoch, 10,000 feet above sea level. An extension of this work to higher altitude would be extremely interesting; perhaps a similar instrument could be taken up in the next stratosphere flight. One should expect a large increase of the effect and an appreciable extension to lower and perhaps also to larger wave-lengths if a height of 60,000 feet or even more could be reached.

[10] Compare I. Bartels, *Elektrische Nachrichtentechnik*, Vol. 10, special number 1933 (*Physik der hohen Atmosphaere*).

[11] The same instrument has been used successfully for measuring the amount of ozone near ground level and the ultraviolet intensity of the sky light (cf. Helv. Phys. Acta **7**, 672, 674 (1934)).

Some Phases of the Mitogenetic Ray Phenomenon

ALEXANDER HOLLAENDER* AND WALTER D. CLAUS,* *Laboratory of Plant Physiology, University of Wisconsin, Madison*

(Received April 17, 1935)

I. INTRODUCTION

"MITOGENETIC rays" are reported to be ultraviolet light of 1900A to 2500A emitted by biological materials in certain stages of development and by chemical reactions, and to have an intensity of 10 to 1000 quanta/cm²/sec. The "radiation" was first reported in 1923 by Professor Alexander Gurwitsch,[19] then of the First State University of Moscow. It was thought originally that the phenomenon was of purely biological significance, and that it was the result, and partially the cause, of cell division. Because of this stimulating influence on mitosis, it was termed "mitogenetic." Now, however, many investigators believe that the emission of this short wave-length ultraviolet light is connected with practically all chemical reactions, and that the source of the radiation in living matter is none other than the reactions taking place within it. For this reason, the radiation is termed by some, "oligoluminescence;"[6] others call it "Gurwitsch rays" and "M-rays."

The mitogenetic ray problem is probably one of the most interesting in modern biology. At the same time, as we come to realize the full significance of the physical characteristics ascribed to it, we find it to be a most difficult and perplexing problem. Many investigators are not at all convinced of the existence of the phenomenon, at least in the form reported, despite the fact that more than 500 papers and a number of books have appeared on the subject. The larger number of publications come from biological laboratories, but a number come also from physical laboratories with excellent international reputations. Most of the recent reports concentrate almost entirely on the use of mitogenetic ray technique in the study of difficult and involved biological phenomena or of chemical problems connected with them. Thus, publications during the past two years treat mostly of the spectral analysis of the radiation arising from biological and chemical reactions. Very little of the published work concerns itself with the methods of detection now in use, for it assumes that such methods have been satisfactorily established.

We have thus a very contradictory situation: Here is one group of investigators working on this most advanced biological phenomenon, sometimes conscious of the apparent physical obstacles to its existence, but often apparently not conscious of the nonconformity of their findings with accepted physical principles. On the other hand, there is a group of workers who, thoroughly convinced of the nonexistence of the rays, believe that the positive results reported by others are caused by faulty technique, by too small a number of experiments or by other factors, and that the phenomenon is one which has nothing at all in common with ultraviolet light. These workers are perhaps supported in their views by the many contradictory findings reported by the investigators who claim positive results. This point of view is very interestingly and capably presented in a recent review by Bateman.[3] For our own part, however, we cannot believe that the phenomenon should be relegated to the limbo of an ignominious past, merely on the ground that some of the evidence presented by possibly overenthusiastic supporters is inconsistent. We hold to the view that the ground work is not yet complete, and that a definite pronouncement is not possible at present.

We will, therefore, in the following discussion, concentrate less on the application of the phenomenon than on the methods of its detection. Many careful workers, well acquainted with the particular material with which they are working, and following in the smallest detail the descriptions given by those reporting successful experiments, have been unable to repeat such experiments. For this reason, we feel that it is inadvisable to discuss applications before the methods of detection have been fully established.

Methods of detection can be divided into two types, biological and physical. The phenomenon was first observed by biological means, and

* Under the auspices of the Radiation Committee of the Division of Biology and Agriculture, National Research Council.

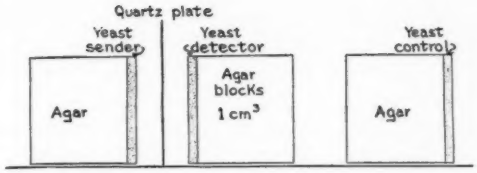


FIG. 1. Diagram of fundamental induction experiment with yeast as both sender and detector (mutoinduction).

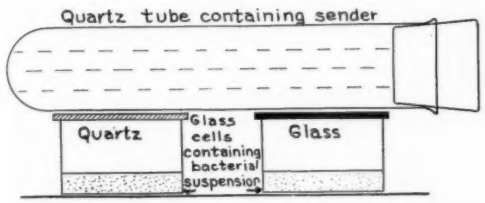


FIG. 2. Diagram of fundamental induction experiment with a chemical or enzymatic reaction as sender, and bacteria as detector.

although successful physical detection was reported some five years ago, the majority of experiments are still carried out with biological detectors.

II. BIOLOGICAL DETECTORS

The first work reported by Gurwitsch[19] was based on the following experiment: single roots of the common kitchen onion, while still attached to the bulb or to a piece of the bulb, were arranged so that one of the roots (detector) enclosed in a glass capillary was held in a vertical position with the growing section near the tip of the root exposed. Another root (sender) was placed in a horizontal position with the tip exposed, the arrangement being such that the projected axis of the sender root struck the detector exactly in the median line. After exposure, the detector root was sectioned and stained, and a preponderance of mitotic figures were found on the side of the root exposed to the sender. (Mitotic figures are characteristic arrangements of nuclear material occurring when the cells are in the process of division.) These results were not substantially different when quartz was placed between sender and detector, but were not obtained when glass was used—hence the opinion that the radiation consisted of ultraviolet light. These experiments have been repeated by a number of workers, some reporting positive results, and quite a large number reporting negative results. Many of the criticisms and doubts of the existence of the radiation are based on the difficulty and weakness of the so-called onion root method. Since detailed discussions of the different points of this method can be found in reviews by Bateman[3] and by Hollaender,[24] they will not be discussed here.

The detectors more generally in use today are yeast and bacteria, for which the arrangement of the fundamental experiment is quite simple, as is shown in Figs. 1 and 2. (A detailed discussion of the preparation of the materials for the detectors is given elsewhere.[24]) Special precautions against chemical effects should be taken. Wolff and Ras,[60] for instance, enclosed the sender in a stoppered quartz tube. The detector also was covered by a quartz plate, and the control by a glass plate. In the Gurwitsch experiment, sender and detector were separated by a quartz plate. The distance between sender and detector may be 0.5 cm to 12 cm, depending on the time of exposure to be given, which may vary from 10 seconds to 60 minutes. The experiments are usually conducted in daylight, and as a matter of fact, yeast is reported to be usable as a detector only in daylight. No attention is usually paid to the possibility that ultraviolet light may be scattered in the laboratory, which in intensity may exceed the possible mitogenetic ray intensity emitted by the sender. The control and exposed material, however, are always handled in an identical manner, except that the control is protected from any possible radiation emitted by the sender. The physical arrangement of these experiments is simple and does not permit of any criticism from the physical point of view. Hence any criticism must arise from the method of handling the biological materials.

The usual procedure is to subject the detector organisms to several minutes irradiation by the sender, to remove a portion of both irradiated and control material to fresh nutrient medium, and to allow this to incubate for several hours, after which the number of organisms per unit volume is determined. The effect shows itself, it is claimed, in an excess in the number of irradiated organisms of twenty to several hundred percent over the controls. Sometimes counts

are made under the microscope, but with bacteria this is a very tedious procedure, since the organisms are usually less than $1\mu^3$ in size. The more usual procedure is to make plate counts: a properly diluted suspension of the bacteria is mixed with melted nutrient agar which is then allowed to cool and solidify in a petri dish which is kept covered to protect the agar from contamination. After a number of hours of incubation, the individual bacteria will have formed visible colonies which are easily counted. In this method, however, not all the bacteria are counted, but only those which have formed separate colonies, since not all the bacteria in a suspension are capable of continued division. It may be assumed that part of the effect of mitogenetic radiation is to stimulate these otherwise dormant organisms to reproductive activity. Other methods[24] also have been used to determine the number of bacteria in mitogenetic ray experiments.

Yeast, until recently, has been used almost exclusively in detection. A large number of experiments have been reported, chiefly by Gurwitsch,[20] in which the increase in the percentage of bud cells of a definite size is used as a criterion of effect. (Most yeasts multiply by the formation of buds which break away from the mother cells at a certain stage of their development.) Other methods determine the *total number* of cells in the exposed culture as compared with the total number in the control culture. Another method compares the *volume* of yeast in the exposed suspension with the volume in the control suspension. All of these methods involve a large number of details, but Gurwitsch has never published a plain clear-cut description with all the details of preparation of the materials, or sufficient description and evaluation of the possible errors involved. As far as descriptions are given, they are often contradictory. It is not unusual for the description of a method of a few years ago to have so changed in essentials during the last year that one cannot understand how any results could have been obtained at all. Although one believes that methods should be improved, one might ask that there at least be consistency in the procedures.

It is just lately, and especially in the laboratories outside of Russia, that yeast has been replaced by bacteria. Bacteria usually multiply by fission, and the value of these micro-organisms as detectors (these remarks apply to yeast also) rests on the characteristic properties of their growth phases. Thus there is little increase in the population of an agar slant culture of *Escherichia* (*Bacterium*) *coli* after about a day. If now, some of these bacteria are transferred to fresh bouillon, the bacteria will not begin to increase in number at once, but will apparently remain dormant for several hours, depending on the age of the culture, the type of the medium from which and to which they are transferred, the size of the inoculum and the temperature. The organisms, however, are not truly dormant during this period (often called the "lag" period) but may grow in size, and toward the end of the period will begin to increase in number at a slow rate. This period of retarded growth is followed by the period of maximum growth, during which time (usually called the "log" phase) the number N of bacteria increase according to the equation

$$N = N_0 2^{(t-t_0)/T},$$

where N_0 is the number at time t_0 when the culture may be considered as entering the log phase, and T is the "generation time" required for one complete cell division. The generation time for yeast is about one hour, but for bacteria under optimum conditions it may be as short as 15 minutes. Thus, one bacterium may multiply 16-fold in one hour. Fig. 3 shows a typical growth curve.*

The differences between the lag and log phases are distinct and definite, and the phases can be predicted with reasonable accuracy if the different factors influencing them have been studied. The mitogenetic effect is reported to express itself in a shortening of the lag phase. Thus, micro-organisms which have been exposed to a mitogenetic ray sender will show an increase in

* The later period during which the number of organisms decreases, sometimes called the "negative growth period," will not be discussed here since it is outside the scope of this article. Good reviews of growth curves, including physiological reasons and mathematical interpretations of them can be found especially in: R. E. Buchanan and E. I. Fulmer, *Physiology and Biochemistry of Bacteria* (Williams and Wilkins, 1928–30). O. Rahn, *Physiology of Bacteria* (P. Blakiston's Son and Co., 1932).

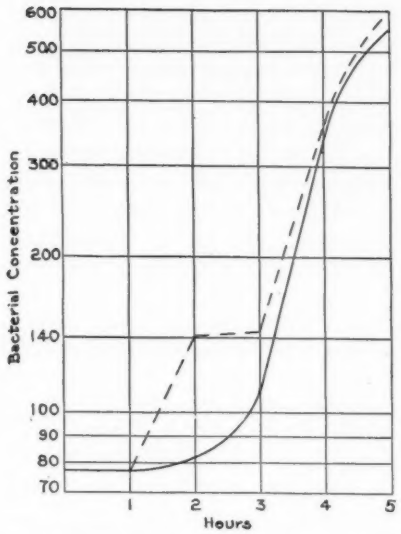


FIG. 3. A typical bacterial growth curve, represented by the solid line. Growth of a bacterial culture after exposure to a mitogenetic ray sender, indicated by broken line. Redrawn from Wolff and Ras.[58]

number somewhat earlier than the controls. This excess does not last, however, and the control will catch up with the exposed culture, and may even pass it. The length of the exposure and the time at which the effect can best be determined must be carefully investigated in each case. For this reason, the best procedure when using bacteria or yeast as the detector is to follow the growth curve of both exposed and control material for several hours after the exposure, as shown in Fig. 3. This method, although more tedious, is much more satisfactory, and is more likely to show the effect if it exists, than if counts were taken only once haphazardly after the exposure.

It has been claimed that these organisms cannot be greatly influenced during the log phase, since they are already enjoying a high growth rate. It is known from other work, however, that they are extremely sensitive and are really most easily influenced during this phase, but it is very difficult to recognize a change in the number of organisms at this time, especially if the change is not very pronounced. For this reason, they are not used as detectors when in the log phase. When they are to be used as senders, however, they should be in the log phase, because the highest emission of radiation is reported to take place during this period of greatest constructive metabolic activity. These points have been especially brought out by Wolff and Ras,[58] and by Tuthill and Rahn.[37, 56] We would like to call the attention of physicists to the possibilities offered by the growth characteristics of micro-organisms as extremely sensitive indicators of physical and chemical factors. It is planned to discuss some of these possibilities at a later date.

All the methods described are reliable within certain limits. These limits can be modified by the care which each investigator takes in handling his materials, in measuring small quantities of liquid, and in many other things with which he becomes acquainted as he tries to use biological materials as an equivalent or as a substitute for physical tools. That this care has been taken is not always obvious from the publications. Through the many possibilities for variation in living materials, effects can too easily appear, especially when too small a number of experiments are conducted or when the number of controls is not sufficient. In most publications on the subject, not enough material is given to judge of this point.

It should be possible to repeat these experiments from published descriptions without too much difficulty, with the possible exception of the method which uses the increase in the percentage of yeast buds as an indicator of effect. The estimation of these buds is a very subjective process, and rests entirely on the judgment of the investigator. All other methods seem more objective. In spite of this fact, a large number of workers, many of whom have apparently done a thorough job, report negative results. If, in this work, special precautions are to be taken, they should be mentioned. We feel that the larger responsibility in this matter rests with the leaders of the work who, too often, have accepted with little criticism the results of investigations based on too few experiments.

III. PHYSICAL DETECTORS

If we assume that the evidence is sufficient to prove the electromagnetic character of mitogenetic radiation, and that the wave-length lies

in the region of the short ultraviolet, i.e., 1900–2500A, then it should be possible, theoretically at least, to detect the radiation by purely physical means. The most obvious approaches to the problem are by way of the photographic plate and the photoelectric cell. Another approach which we shall consider first, depends on the ability of light to disturb certain delicate chemical reactions, of which the phenomenon of Liesegang rings is an interesting example. When a thin layer of gelatin containing ammonium bichromate is spread over a glass plate and a drop of silver nitrate solution is placed in the center, a number of concentric rings of silver chromate are precipitated as the silver nitrate diffuses radially from the drop. If now, a short distance above such a system, there is supported a plate having a slit which extends across a diameter of the rings (about 5 cm), the symmetrical development of the rings will be easily disturbed by the influence of light or vapors coming through the slit. Cellophane or quartz is placed over the slit and the inductor (onion pulp, carcinoma, etc.) placed thereon. Stempell,[52] and Stempell and Romberg,[53] have published a large number of photographs showing distorted ring formation (the "Stempell effect") along the direction of the slit. These distortions, the authors conclude, are the result of mitogenetic rays emitted by the biological inductors. Siebert,[50] and others, although able to repeat many of these experiments, maintain that the effects through cellophane are due entirely to chemical influences, while experiments with quartz covering the slit are inconclusive.

In other experiments, Stempell[52] exposed droplets of H_2O_2 to germinating peas, onion roots and other biological senders, and reported that the H_2O_2 decomposed more quickly, as determined by colorimetric reactions, when exposed to the senders. Mardaschew and Mogliewsky[34] noticed that when the two enzymatic systems, arginin-arginase and saccharose-saccharase, were exposed through quartz to blood and other biological senders, there was a slight reduction (maximum 26 percent) of the enzymatic activity. Their experiments with Liesegang rings and other chemical detectors are, however, not very impressive.

The photographic situation is just as unsatisfactory. A great many experiments, positive and negative, have been reported, of which we shall mention only a few. Rajewsky[38] calls attention to the work of Eggert and Noddak[12] who found that an energy of at least 9×10^{-3} ergs/cm^2 is required to give a just perceptible blackening on a photographic plate. But since the intensity of mitogenetic radiation is so small, and since the failure of the reciprocity law is greatly accentuated for very long exposures, he calculates that an exposure of at least 92 days would be required to give a perceptible photographic image.

Loos[31] exposed ultra-rapid plates, unsensitized and sensitized with paraffin oil and with vaseline, through quartz, to onion pulp which he changed 3–5 times daily for 8–10 days. He also exposed his plates to a "mitotin-mitotase" mixture, changed 4–6 times daily for 9 days. He obtained no images. Rossmann[43] obtained negative results using both orthochromatic and Schumann plates. Taylor and Harvey[55] exposed Kodak Superspeed film through 0.2 mm of quartz to growing onion roots for 24 hours, and to yeast cultures renewed at intervals of 2–5 days for as long as 89 days, without effect. Lorenz[32] exposed Eastman Speedway plates, unsensitized and sensitized by Nujol, through quartz, to onion pulp and onion root tips for as long as 184 hours without effect. In earlier experiments by one of us (A. H.), high speed plates, unsensitized and sensitized, were exposed in a specially constructed, light tight box to the usual biological senders (onion pulp, yeast, etc.). Although exposures were extended over a period of several months, and the senders were renewed daily, no effect on the plates could be observed.

On the other hand, Protti[36] introduced strips of film 1.5×20 mm, enclosed in quartz, into the peritoneum, muscular and subcutaneous tissue of rabbits, into the muscles of eels, and into onions and cactus leaves, and obtained photographic effects of various intensities! Reiter and Gabor[41] conducted a number of experiments with high sensitivity and Schumann plates which gave negative results when carried out in the dark. They came to the conclusion that if mitogenetic radiation was to be detected photographically, it would have to be done while the senders were exposed to light. They devised a method whereby onion pulp was illuminated by weak daylight

filtered through 5 cm or more of paraffin oil to remove all blue and ultraviolet. Between the pulp and the photographic plate was interposed a filter which would transmit some long wave-length ultraviolet (2600–3500A?) and only red in the visible region. When white paper was substituted for pulp, two days' exposure was required to give a minimum blackening on the plate. The box was then divided into two halves with narcotized onion pulp on one side and fresh pulp on the other, renewed hourly for two days. These experiments showed a decided difference in photographic effect on the two sides. Only two such experiments were performed, however, and authors themselves are not entirely convinced of the conclusiveness of the work. Note that these experiments support the contention of these authors that the wave-length of mitogenetic rays is that of the longer ultraviolet.

In connection with the experiments of Reiter and Gabor, one should recall that not only is a minimum *energy* required to give a photographic image, but also a minimum *intensity* below which no prolongation of exposure time will produce an effect. It is quite possible, however, that a subminimal intensity superimposed on a greater-than-minimal intensity would produce an observable effect in reasonable time. This is substantially what Reiter and Gabor did, although they superimposed presumably ultraviolet light on long wave-length visible light. Interesting questions arise in this connection, but we are not aware at present whether they have been investigated or not. It is generally conceded, however, that the photographic plate is a very unsatisfactory detector of extremely weak intensities.

It therefore appears that if our *a priori* proposition: that extremely weak ultraviolet radiation can be detected by physical means, is to receive experimental verification, it must be with the help of the photoelectric effect. Unfortunately, the "photoelectric yield," or number of ejected photoelectrons divided by the number of incident quanta, is far from being a 1 : 1 ratio. It is generally assumed[4, 18] that a good average yield in the ultraviolet is $1 : 10^3$. This opinion appears to be based mainly on the work of Suhrmann[54] on platinum, gold and silver foils, wherein he found that the yield of platinum and silver *under optimum conditions* is of the order of $1 : 10^3$, although his best yield for gold was about ten times less. As we shall see, this assumption is not to be taken too literally. Actually, of course, the yield rises from zero at the threshold to one or more maxima, the height and location of which are characteristic of the surface used, and for the common metals under ordinary conditions, is apt to be between $1 : 10^4$ and $1 : 10^6$ in the ultraviolet. Greater yields have been recorded,[26] but they are ordinarily available only with visible light incident on specially treated sodium and potassium surfaces.

If now we consider the intensity of mitogenetic radiation as being of the estimated intensity of 10 to 1000 quanta/cm^2/sec., we could hardly expect from the average photo-cell more than one photoelectron/sec., the detection of which depends on the amplifying device used. But the minimum current detectable by DuBridge,[9] using a balanced FP–54 circuit under very favorable circumstances, was 32 electrons/sec.!

If, however, gas be admitted into the photocell and a voltage just below the sparking potential applied between the cathode surface and the anode loop, the ejected photoelectrons are drawn through the gas with sufficient speed to produce ions by collision, thus giving enough current to be easily detected. As Elster and Geitel[11] found, however, such a current is subject to violent fluctuations for weak incident intensities. Schreiber and Friedrich,[46] working with such a device, found that it was feasible to convey the impulse created by a single high speed photoelectron either to a fiber electrometer or through a small capacitance to an amplifier and mechanical register. A resistance of about 10^{10} ohms connected the anode to ground, and permitted the charge to leak off fast enough to prepare the cell for the next discharge. In this manner, Schreiber and Friedrich were able to count *individual* photoelectrons as they were ejected, without having to depend on their cumulative effect to produce a measurable current.

Shortly before the announcement of this work appeared, Rajewsky[38] had published an account of his special photoelectric modification of the Geiger-Müller[15] ion counter which had been in use for several years to count α- and β-rays.

(Because of this priority in publication, and because his device is slightly more sensitive, Rajewsky is usually given credit for devising the photoelectron counter. In a later paper, Schreiber[47] takes particular care to point out that their work was actually simultaneous and independent.) The modified Geiger-Müller tube consists primarily of a cylinder of photosensitive material 1–2 cm in diameter and 2 or 3 times as long, with a fine wire coaxial with the cylinder and highly insulated from it. Gas is admitted to a pressure of several cm, and a high potential is applied between the cylinder as cathode and the wire as anode. For such an arrangement the electric potential increases logarithmically toward the axis. Light entering a window in the cylinder falls on the opposite wall and ejects a photoelectron which is drawn toward the wire, creating a burst of ionization in its immediate vicinity. Connections to ground and to a registering device are similar to those of Schreiber and Friedrich, a diagram of which is shown in Fig. 4. Detailed descriptions of further designs of tubes and registers are given by Rajewsky[39] and Locher.[30] A constant voltage source which we have found very satisfactory is described by Schmitt.[45] Other high voltage and amplifier circuits are described by Dunning.[10]

Unfortunately, it is not possible to detect weak intensities with a Geiger-Müller counter merely by placing the suspected source in front of the counter window and waiting for impulses. Even when the tube is encased in several cm of lead, stray γ-rays, cosmic rays and spurious tube discharges will give a "dark" count varying from a few to 50 or 100/minute. Thus it is necessary to alternate dark and exposed counts for periods of 5 to 15 minutes each. Now the mean statistical deviation for such random counts is equal to the square root of the *total* number of counts, N, so that if a positive effect is to be claimed, the difference between dark and exposed counts must be at least $N^{\frac{1}{2}}$. But if a real effect exists, no matter how small, the difference may be brought to any arbitrary, predetermined value merely by taking a sufficiently large number of counts. 3 to $5\times(N)^{\frac{1}{2}}$ is usually considered to be a convincing demonstration.

Further complications arise in the use of the Geiger-Müller tube, in that it is essentially an

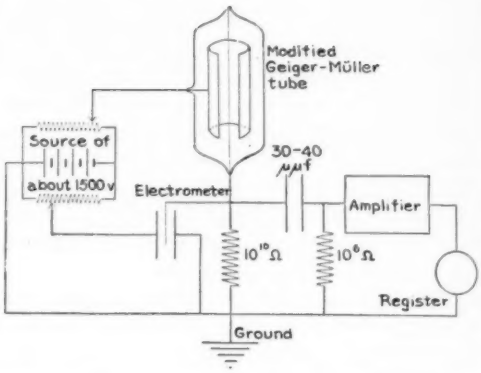


FIG. 4. Diagram of modified Geiger-Müller counter with simplified accessory circuits.

electrostatic device which must be thoroughly protected by a grounded metal shield. Grounding its lead casing is usually sufficient. Leads from the high voltage device and to the amplifier must also be shielded. The tube must not be exposed to high humidity, as a film of moisture between electrodes may permit slight fluctuations in leakage current which will be amplified and registered. Gray and Ouellet[17] record that onion effluvium (!) exerts a distinct influence on the count. One must therefore be exceedingly careful about placing moist material on or near the tube. Electrical discharges such as are sometimes used to tetanize muscle preparations may introduce spurious discharges which may be misinterpreted as a mitogenetic ray effect. Occasionally an author will describe in detail how he eliminated such sources of error, but too often the reader must accept a casual statement that such possibilities were removed, or they are ignored entirely. Still greater and more subtle chances for error creep in during the determination of the sensitivity of the Geiger-Müller tube, so that it is small wonder that publications on this type of work report such widely varying results. Table I gives a summary of the available published reports of work in which the Geiger-Müller tube was used as a mitogenetic ray detector.

An examination of the table reveals the curious fact that six groups of workers profess to have obtained positive evidence of mitogenetic rays from biological preparations, or their counterpart from chemical reactions *in vitro*, while six groups

TABLE I.

AUTHOR	TYPE OF CELL	SENSITIVE SURFACE	SENSITIVITY	WAVE-LENGTH	METHOD OF CALIBRATION	MATERIALS TESTED	RESULT	ESTIMATED INTENSITY
Rajewsky 1930–31	Modified G-M Tube	Cd	$1:10^2$ to $1:10^3$	2650A	Cellophane filters	Onion roots Onion pulp Carcinomas	+ + +	10–100 quanta/cm²/sec.
Schreiber and Friedrich 1930	Modified Photo-cell	K	$1:10^4$	3660A 2660A	Photographic emulsions, Crossed Nicols	Yeast	–	<2000 quanta/sec.
Frank and Rodionov 1932	Modified G-M Tube and Photo-cell	Al Cd	$1:6\times10^3$	2537A	Solution filter of variable thickness	Frog muscle Frog muscle pulp Frog hearts Inorg. chem. reactions	+ + + +	100–1000 quanta/cm²/sec.
Locher 1932	Modified G-M Tube					Onion roots Bacteria	– –	
Seyfert 1932	Modified G-M Tube	Zn-amal.	$1:5\times10^2$ $1:10^5$	2300A 3400A	Calculated from standard lamp at 260 meters	Onion pulp Onion roots Yeast	– – –	<5 quanta/cm²/sec.
		Al Mg	$1:3\times10^3$ $1:5\times10^5$	2300A 2300A		Chick embryos Mouse tumors	– –	<30 quanta/cm²/sec. <50 quanta/cm²/sec.
Audubert and Doormaal 1933	Modified G-M Tube	CuI CuO Cu_2O Cu Al	Not constant			Chemical oxidations in vitro	+	
Siebert and Seffert 1933	Modified G-M Tube	Cd				Blood Carcinomas Urine Oxidation models	+ + + +	
Gray and Ouellet 1933	Modified G-M Tube	Pt	$1:6\times10^3$	2500A	Calibrated slits	Sea urchin eggs Spermatozoa Yeast Oxidations	– – – –	<50 quanta/cm²/sec.
Petri 1933	Photo-cell	Cd				Sprouting wheat	+	
Lorenz 1934	Modified G-M Tube	Cd	$1:1400$ $1:2400$	2300A 2537A	Filter of $K_2Cr_2O_7$ sol.	Onion pulp and roots Mouse sarcoma Mouse embryo Frog muscle	– – – –	<15 quanta/cm²/sec.
Barth 1934	Modified G-M Tube	Al Cu CuI				$K_2Cr_2O_7+FeSO_4$ Digestive reactions	+ +	
Kreuchen and Bateman 1934	Modified G-M Tube	Cd Al Zn	$1:2–5\times10^4$	2537A	Filters of ground quartz, Multiple reflections from polished quartz surfaces	Yeast Frog muscles	– –	<300 quanta/cm²/sec.
Hausser and Kreuchen 1934	Modified G-M Tube	Zn Zn (sens.) Cd Fe, Al, Cu Brass	$\{1:10^7$ $\{1:10^{13}$ $\{1:10^5$ $\{1:10^{13}$ $\{1:10^6$ $\{1:10^{13}$	2500A 3660 2500 3660 2500 3660	Filters of ground quartz, Multiple reflections from polished quartz surfaces			
Karev and Rodionov 1934	Modified Photo-cell and G-M Tube	Al Al	$1:10^5$ $1:10^9$	2537A 2537A	Quinine solution of variable thickness			
Reiss 1935	Modified G-M Tube	Cd	$1:10^5$ $1:4\times10^{10}$	2560A 3350	Reflection from quartz, Glass filters			
Suhrmann 1925	Photo-cell	Pt (heated 15 sec.) Ag (outgassed) Au (outgassed)	$1:1270$ $1:3190$ $1:3\times10^3$ $1:11\times10^3$	2240A 2540 2540 2240				

deny the existence of such radiation, at least in detectable amounts. Neither is the evidence presented by either side very convincing when one considers the numerical data obtained. Only two of the six proponents carried through a calibration of their counters whereby they could give an estimate of the intensity they observed, these two placing it somewhere between 10 and 1000 quanta/cm²/sec. The opponents were of necessity more thorough, all placing an upper limit on the possible intensity somewhere between 5 and 300 quanta/cm²/sec. None of the

eight sensitivity determinations, however, is entirely free of question. Our own experience in the use of filters as a means of reducing a measurable light intensity (*ca.* 1 erg/sec., or 1.25×10^{11} quanta/sec. at λ2300A) to intensities which a Geiger-Müller counter can handle (10^3–10^5 quanta/sec.) reveals a tremendous possibility for error.

For monochromatic light, Beer's law states that

$$I = I_0 10^{-\epsilon cd},$$

from which the molecular extinction coefficient ϵ for a solute may be written

$$\epsilon = (1/cd)\ \log_{10}\ (I_0/I).$$

I_0 is the intensity as measured after the light passes through an absorption cell having plane parallel sides d cm apart and filled with the pure solvent. I is the intensity as measured after the light has passed through the d cm of solution of molar concentration c. For purposes of illustration, let us quote from one of our own experiments. We used a Bausch and Lomb monochromator set for λ2300A, slit widths 0.5 mm, and a cell having quartz sides 1.58 cm apart. For K_2CrO_4 in concentrations between 0 and 1.8×10^{-3} M, ϵ could be measured by a thermopile with light from a quartz capillary mercury arc; or by the Geiger-Müller counter using monochromatic light from a microburner gas flame adjusted to give counts of 100/min. or less. Log I plotted against c gave an excellent straight line, from the slope of which ϵ could be calculated. But when the mercury arc was used as a source, and sufficient K_2CrO_4 (4.7×10^{-3}–5.7×10^{-3} M) used to reduce the intensity to within the operating range of the counter, ϵ was found to be 12 percent smaller. Either value will fit published curves satisfactorily in this region, but the difference is sufficient to affect the calculated sensitivity of the tube by a factor of approximately ten. If the slits of the monochromator are widened to obtain more intensity, the difference in the values of ϵ as measured by the two methods becomes still greater, with a corresponding increase in error in the determination of the tube sensitivity. What is the cause of these discrepancies? Experimental error is too small to account for the differences in the values of ϵ. It is not a matter of counter response, because the same range of intensities was incident on the counter in both cases, i.e., sufficient to operate the counter between a few and 100 counts/min. Beer's law may, we believe, be held valid for the range of filter concentrations used: 0 to 5.7×10^{-3} M. The best explanation appears to be that, although Beer's law is valid for *monochromatic* light, the light coming through the monochromator is far from being monochromatic *as seen by the counter*, which varies tremendously in its sensitivity to different wave-lengths. It is of interest to note[22] that for practically all metals, the photoelectric yield increases about 10^6 times as the wave-length in this region decreases by 500A, so that extraneous wave-lengths in the test beam will not behave according to calculations. We attempted to avoid this difficulty by using as a filter a solution of K_2CrO_4 for which the extinction coefficient reaches a minimum at our test wave-length of 2300A and increases rapidly on both sides. This means that extraneous wave-lengths will be absorbed first, which probably accounts for the fact that our value of ϵ at low filter concentrations is greater than at high concentrations. Probably the latter value is closer to correct than the former, but unless one is able to obtain the same value of ϵ at both ends of the intensity-concentration curve, he must be extremely careful about drawing conclusions.

The paper of Kreuchen and Bateman[29] also calls attention to this question, and propounds the following theses: first, that the sensitivities of Geiger-Müller tubes have been greatly overestimated because of failure to consider the presence in the test light, of scattered quanta having a greater photoelectric effect than the test quanta, which effect is greatly accentuated by the use of filters; second, that the photoelectric yield of a counter tube may approach but not exceed the yield of the same surface when used as a photo-cell; third, that the photoelectric yield can hardly be expected to be increased above $1:10^3$, and hence it is manifestly impossible to hope ever to detect mitogenetic radiation by this means.

As to the first point, we have already indicated to some extent our own experience. Let us look briefly at some of the methods used to reduce an

intensity which is measurable with a thermopile, to one which will operate a counter. Rajewsky[38] used a monochromator to obtain λ2650A having an intensity of 112 ergs/cm²/sec., and reduced this to about 12 quanta/cm²/sec., the reduction factor of 1.23×10^{12} being effected by inserting sheets of cellophane in the test beam. Presumably the absorption factor for each sheet was determined separately by means of a thermopile. Such measurements would give a reasonably accurate value of the absorption coefficient of cellophane *for λ2650A*. An intensity of 112 ergs/cm²/sec., however, implies a fairly broad band of wave-lengths, of which those longer than 2650A will penetrate the *complete* cellophane filter in greater proportion than calculated for 2650A, while those shorter than 2650A, although more highly absorbed than calculated, will be more efficacious in activating the counter. These same considerations hold in the use of solution filters, and one is inclined to doubt the high values of sensitivity reported by both Rajewsky and Lorenz.[32] Some workers[28, 14] eliminate the question of the validity of Beer's law by making use of Lambert's law—that is, they use a filter consisting of variable thicknesses of a solution of constant concentration. The method is commendable for this reason, but it does not eliminate the question of extraneous wave-lengths.

Schreiber and Friedrich[46] used a filter of blackened photographic emulsion and crossed Nicols to obtain a reduction factor of 10^8, but they give no details as to the method by which they arrived at the numerical value of this factor. Seyfert[49] calculates the energy between 2700A and 3130A available from a Hefner standard candle at a distance of 260 meters, and observed the response of his counter at that distance; then using Uspensky's[57] method, he calculates the sensitivity of the counter at 2300A and at 3400A. Superficially, it would seem that there are too many possibilities for error to place very much confidence in his value of 5 quanta/cm²/sec. which he claims he could detect.

The work of Hausser and Kreuchen,[22] and Kreuchen and Bateman[29] appears to have been carefully done. They used filters of ground quartz plates, combined with multiple reflections from polished quartz surfaces. The proportion of intensity reflected at each surface could be calculated from optical laws. Here again, however, the validity of the method would seem to depend somewhat on the degree of monochromaticity of the original beam. They used an original intensity of approximately 10^{13} quanta/sec. from a double monochromator.

Perhaps the least objectionable method for reducing intensity for use with a Geiger-Müller tube is that of Reiss,[40] who used glass filters of various kinds for the reduction of mercury arc wave-lengths down to about 3000A, and reflection from quartz plates for shorter wave-lengths. To eliminate the effects of extraneous light, he determined the dark count by turning into the counter the regions of the spectrum immediately adjacent to his test line, and assumed that the average of counts on each side equaled the count due to scattered light in the test line. Unfortunately, he worked down only to 2560A, and there is some doubt as to the practicability of the method for the lines of weaker intensity which lie below 2536A.

Numerical values for all these determinations will be found in Table I.

The second point made by Kreuchen and Bateman,[29] that the photoelectric yield of a surface in a counter tube may approach, but not exceed, the yield of the same surface in a vacuum photo-cell, would appear to be more or less valid, although it is well known that the presence of gases and water vapor will sometimes greatly affect the photoelectric yield.[27] For this reason, it does not necessarily follow that a yield as determined in vacuum and with high incident light intensity will be equal to the yield of that same surface in a counter tube. This is unfortunate, as the yield in vacuum may be determined with reasonable accuracy. Peculiarly enough, however, the presence of gases, especially of hydrogen and water vapor, usually increases the photoelectric yield, while as Kreuchen and Bateman point out for cadmium at λ2536A, the yield as a counter is only 0.01 of the usual photoelectric yield. Karev and Rodionov[28] made actual determinations with the same aluminum surface. They used it first as a photo-cell operated at the optimum voltage of 300 v, then admitted 1.8 mm of hydrogen and used it as a Geiger-Müller tube at 480 v, and found that the yield

as a counter was about 0.001 of the yield as a photo-cell over a considerable range of wave-lengths. Such inequalities are very perplexing, but there are two possible explanations: (1) All photoelectrons ejected by the counter are not counted, in which case better centering of the anode wire and polishing of the surface to give a more uniform distribution of the electric field might help; also a higher gas pressure with the corresponding increase in operating voltage will give a greater potential gradient at the surface, and will definitely assist in garnering photo-electrons; (2) all photoelectrons may be registered, but possibly there is a change of surface conditions on discharge, with so slow a return to normal, if at all, that only a part of the incident light is effective.

The third point of Kreuchen and Bateman;[29] that the photoelectric yield can hardly be increased above $1:10^3$, and hence there is no hope of detecting mitogenetic radiation by this means, is, it seems to us, unduly pessimistic. The photoelectric yields in the visible region for sodium and potassium have been increased tremendously by special treatment of the surfaces.[26] Sensitization by a glow discharge in hydrogen has increased the efficiency of zinc and copper surfaces to ultraviolet many fold, and several investigators[27] have found the halides, especially the iodide of copper, to have a higher yield than the pure metal. Experiments at the Westinghouse laboratories[48] have shown thorium and uranium to be especially sensitive to ultraviolet light. Curves published by Young and Pierce[63] show a rapidly increasing yield for a CsO surface in the region of the short ultraviolet. It would seem, therefore, that there is small reason to despair so soon about the problem of detecting weak ultraviolet intensities.

Let us mention briefly at this point the question of the Geiger-Müller tube as a quantitative instrument. Audubert and Doormaal remark that the sensitivity is a function of so many things that it cannot be determined. Barth[2] observed an extremely rapid fatigue of his surfaces, while Karev and Rodionov[28] present extensive data to show that the sensitivity of their counter increased about 5 times as they slowed down the counting rate from 90/min. to 3/min., and give a formula for the yield as a function of the intensity. Since, however, these data are apparently based on the use of a quinine solution filter, the extinction coefficient of which was obtained at high light intensities and used at low intensities as per the discussion above, we are not inclined to give the report too much weight. In all our own work, log counts plotted against the concentration of K_2CrO_4 used as a filter resulted in exceedingly good straight lines, providing we kept the counting rate at 100/min. or less (the same range as that used by Karev and Rodionov). Locher[30] reports that the discharge rate of his counters varied inversely as the square of the distance from the light source. These two results could only be obtained if the Geiger-Müller counter were a strictly quantitative instrument.

A summary of the attempts to detect mitogenetic radiation with photoelectric devices follows. Rajewsky[38] (F. Dessauer's laboratory, Frankfurt a. M.), using onion sole and roots, and carcinoma cells, obtained counts 20–30 percent greater than his average dark count of 40/min. From this, one estimates a mitogenetic ray intensity of 10–100 quanta/cm²/sec. Frank and Rodionov[14] (A. F. Joffé's laboratory, Leningrad) obtained 30–60 percent increases in count from frog muscles, frog hearts and muscle pulp. With pyragallol + NaOH, a 50 percent increase over the dark count was observed. The reaction of $SnCl_2$ + $HgCl_2$ increased the count by 13 percent. $K_2Cr_2O_7 + FeSO_4$, reacting in the dark gave an increase of only 9 percent, but in visible light it resulted in an increase of 30–60 percent. Needless to say, the visible light was screened from the counter with a monochromator. Their data place the intensity of mitogenetic radiation at 100–1000 quanta/cm²/sec. Audubert and Doormaal,[1] using a number of oxidation reactions, obtained counts from 2 to 14 times the dark count! They specify, however, that the tube sensitivity is not constant, and that the phenomenon of fatigue requires frequent renewal of the cathode surface. Moreover, they claim the spectral emission is characteristic of the reaction, and the proper cathode surface should be used to detect it: CuI for the extreme ultraviolet, Al between 2000A and 2200A, and Cu_2O and CuO for wave-lengths closer to the blue-violet.

Siebert and Seffert[51] made simultaneous use of

two similar cadmium counters and showed that the variation in the dark rate between the two was never greater than 3 percent. When blood, carcinoma cells, urine and certain oxidation models held on filter paper were placed before one of the counters, the difference in counting rate rose to 40–50 percent of the dark rate. As this difference disappeared when the exposed blood cells were treated with KCN, it would seem that it was not due to increased humidity or electrical disturbances. Petri[35] used a cadmium surface photo-cell connected to a micro-electroscope of high insulation. Electrodes and electrical connections were tightly enclosed together with a drying agent. The rate of discharge of the electroscope was observed over 3–4 hour periods alternately with and without 100 germinating wheat kernels before the cell window. The presence of the germinating wheat resulted in a distinctly greater rate of discharge.

Barth[2] (W. Gerlach's laboratory, Munich) gives extended data on several experiments. Using an aluminum surface, he obtained a 53 percent increase in count with the reaction $K_2Cr_2O_7 + FeSO_4$ *in the dark* (see report of Frank and Rodionov). Using a copper surface, he obtained the following results: egg white+gastric juice, 21.6 percent increase; fibrin+pepsin+HCl, 15 percent; fibrin+trypsin+Na_2CO_3, 27.9 percent. As a control experiment, counts were taken on water behind glass and quartz, with only a 2 percent difference in counts. With a CuI surface, fibrin+trypsin+$NaHCO_3$ gave a 38.8 percent increase in count. All these differences, with the exception of the control experiment, were greater than $3\times(N)^{\frac{1}{2}}$, the criterion for effect as discussed earlier in this section. By alternating dark and exposed counts every two minutes, Barth was able to show how the emitted radiation intensity mounted to a maximum in 2 or 3 minutes, then gradually decreased over a period of about 20 minutes. All of his surfaces, however, exhibited strong fatigue. Copper surfaces had to be freshly sanded before each day's work, and had to be discarded after a week. A CuI surface was good for a week without further treatment. Barth's published reports appear to give every indication of dependability, and in a personal communication he states that he has been able to repeat similar experiments at will. Sensitivity determinations have been delayed, unfortunately, but are to be carried through as soon as possible.

As against this array of positive reports, we have an equal number of failures to detect mitogenetic radiation with Geiger counters. Schreiber and Friedrich[46] could obtain no extra counts from yeast, and placed an upper limit on the emitted radiation, if it existed, at 2000 quanta/sec. Locher,[30] although not primarily interested in the problem of mitogenetic radiation, placed onion roots and bacteria before his counters, and was unable to observe an effect. Seyfert[49] investigated onion roots and pulp, yeast, chick embryos and mouse tumors with Geiger-Müller tubes having surfaces of magnesium, aluminum and zinc-amalgam. He observed no increase over the dark count, and was able to place an upper limit on the mitogenetic ray intensity at 5 quanta/cm²/sec. Gray and Ouellet[17] constructed a Geiger-Müller tube by coating the inside of a quartz tube 11 cm long and 1.5 cm diameter with a translucent layer of platinum. Fertilized sea-urchin eggs were sealed in another quartz tube held parallel to the counter, and particular care was taken to make observations in all stages of development. No radiation was detected and since the authors claimed that their counter could detect 50 quanta/cm²/sec., the eggs could not have emitted more than one quantum every 40 seconds. To test emission as a function of metabolic rate, eggs in the blastula stage, and spermatozoa, were exposed to the counter without positive effect. Yeast cultures in quartz petri dishes were exposed at intervals over a period of 72 hours, while certain chemical reactions sealed in a quartz tube were tried, all without positive result. The authors take special care to point out that if the counter is exposed to vapors and volatile substances, serious disturbances in counting rate will result.

Lorenz,[32] using a cadmium surface Geiger-Müller tube for which he claimed high sensitivity, could obtain no increase in count from onion pulp and root tips, mouse sarcoma, mouse embryo tissue and tetanized from muscle. He placed an upper limit on the intensity of the emitted radiation, if any, at 5 quanta/cm²/sec. Kreuchen and Bateman,[29] using surfaces of

cadmium, aluminum and zinc, could obtain no evidence of radiation from yeast and stimulated muscle when scattered light, electrical effects and moisture were excluded.

One feels somewhat baffled by such conflicting evidence, especially as there is what appears to be reliable work on both sides. It is our present opinion, however, that negative results do not necessarily carry the significance usually attached to them, because of over-estimated counter sensitivities. The reports of positive effects are less easy to account for, but we believe that until adequate sensitivity determinations are available in these cases, it is best to hold conclusions in abeyance. The present status of the problem of physical detection of mitogenetic radiation is, to say the least, extremely unsatisfactory.

IV. Properties of Mitogenetic Rays

1. Secondary radiation

How is it that organisms respond to such extremely small intensities of ultraviolet light when they are exposed in a nutrient medium which is highly absorbing for these wave-lengths? This is one of the most difficult phases of the mitogenetic ray problem, and an understanding of it is directly necessary to a full understanding of the phenomenon. It brought Gurwitsch[20] to the concept of "secondary radiation," a concept hard to follow and very hazy, even in the description given by him. When a cell is stimulated by a quantum of ultraviolet light, it is supposed to give up light of the same or different wave-length in a quantity larger than it originally received. Also the medium in which the cells are suspended, though highly absorbent for ultraviolet light, transmits mitogenetic radiation by some means such as chain reactions, perhaps. The passage of the radiation through this highly absorbing material is not conditioned by the presence of the organisms, for suspension material from which the organisms have been entirely removed will also act as a secondary radiator. Experiments described by Wolff and Ras[59] (see Fig. 5) show that the intensity of the secondary radiation is greater than that of the incident radiation, for exposure through secondary radiators will give effects in less time than direct exposures. Such secondary radiators become fatigued after 3 or 4 hours use, however, but recover if allowed to rest for several days. A number of other features of this secondary radiation are described by Wolff and Ras.

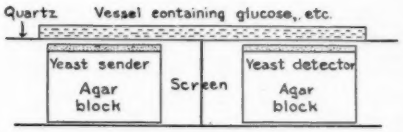


Fig. 5. Diagram illustrating mitogenetic induction through a secondary radiator.

2. Wave-length determinations

The wave-length now generally accepted as that of mitogenetic rays by investigators reporting positive results is between 1900A and 2500A, although Reiter and Gabor[41] reported that their experiments showed a maximum at 3400A and another much smaller one at 2800A. These determinations are based on the effect of artificial light which was passed through a light-strong quartz spectrograph. Onion roots replaced the usual photographic plate, and a preponderance of mitotic figures was found to have developed on the side of the root turned toward the prism, at the places where the wave-lengths 3400A and 2800A would appear on the photographic plate. Their tests for wave-lengths below 2500A gave absolutely negative results. They found also that filters which would exclude light below 2600A did not exclude positive effects. No explanation has been given for the discrepancy between the wave-lengths found by Reiter and Gabor and by the Gurwitsch school, except that they may be dealing with entirely different phenomena.

The wave-length determinations of the Gurwitsch school rest on the following bases:

1. Absorption of mitogenetic radiation by glass and not by quartz; also by all filters which do not transmit below 2600A.

2. Spectral dispersion of the light supposedly emitted by biological senders and chemical reactions, and the effect of this light on the biological detectors which take the place of the photographic plate at the spectrograph.

3. The stimulative effects on biological materials of artificial light having a wave-length less than 2600A.

That mitogenetic rays are absorbed by glass and not by quartz has been reported by most investigators observing positive effects. As a

matter of fact it is customary to protect the control against mitogenetic rays by means of a glass plate. A number of experiments[58] have been reported with other filters, such as 0.5 percent KNO_3 and 0.02 percent $Mg(NO_3)_2$, which have cut-offs at 2550A and 2450A, respectively.

The most direct method for wave-length determinations is the spectral analysis of the radiation emitted by the mitogenetic ray senders. There are, however, so many difficult and doubtful points involved that this work must be accepted with great caution, even if the *existence* of the rays should become well established. As this represents the most advanced work on the subject, it will be discussed further on.

The stimulative effect of artificial light on biological materials used as mitogenetic ray detectors has been reported by Chariton, Frank and Kannegiesser,[7] also by Ruyssen[44] and others. Only monochromatic light below 2600A gave the stimulative effect, but Chariton, Frank and Kannegiesser found that about 10^8 quanta/cm^2 were required, whereas a photoelectric cell sensitive to one-third of this amount of energy was not sensitive enough to detect any radiation emitted by biological senders. These authors explain that, although artificial monochromatic light will stimulate if applied in sufficient amounts, the particular wave-length combinations emitted by biological senders are much more efficient.

Work on the stimulative effect of short ultraviolet light needs a thorough verification. It is a very important problem aside from its significance in the detection of mitogenetic rays. Several workers[8, 13, 25, 62] have found that between 10^{-5} and 10^{-6} erg is necessary to inactivate one bacterium. One might suspect that somewhat less energy would have certain physiological effects not necessarily showing themselves as lethal. If these effects were stimulatory, it would not be surprising, but to recognize them, extensive tests must be conducted. Reports[7, 44] of large stimulatory effects will not be convincing until they have been thoroughly tested, and the reports giving the facts are so detailed that the work can be repeated at will. A number of articles have appeared in which the authors[42, 49] report negative results, although they used approximately the same physical conditions and light intensities which gave positive effects in the hands of other workers. This is not surprising in view of the present status of the problem.

There are several theoretical doubts concerning the possibility of the wave-lengths being below 2600A, from the following considerations. It is well known that for all biochemical reactions, the energy available per molecule of reacting substances is insufficient to produce a quantum of ultraviolet light, and this fact has been used to argue the impossibility of ultraviolet or mitogenetic radiation arising from such reactions. Such an argument, however, is based entirely on *average* energies of reaction. It may happen that an occasional molecule will come into an activated state just at the moment of reaction, and give rise to a release of energy considerably greater than the average. Such a possibility would account not only for the energy available, but also for the spectral distribution which, it is claimed, is characteristic of the particular reaction. Unfortunately, mitogenetic ray technique is hardly sufficiently developed at present to permit a quantitative test of such a hypothesis. (See Bateman[3] for a somewhat more extended discussion of ultraviolet chemiluminescence in liquid systems.)

3. Quantitative effects

Another peculiarity of mitogenetic radiation is that it does not affect biological detectors quantitatively. Thus, a greater intensity will not give a greater effect, although it will permit the time of exposure to be reduced, or the distance between sender and detector to be increased. One might deduce from this that the effect is of a quantized nature, and that only a few quanta are required to produce the maximum effect. Such an idea, however, is hard to reconcile with the fact which we discussed under the technique of biological detectors: that experimental material is sometimes exposed to scattered ultraviolet light much greater in intensity than the mitogenetic ray intensity. Effects do not always show themselves by increased growth, but also occasionally by retarded growth.[20] Overexposure, it is claimed, will give retarded growth which may be followed by another period of increased growth.

V. Senders

Mitogenetic radiation is supposedly emitted by a great number of materials. One might even call it a universal phenomenon, provided that convincing proof of its existence could be brought. But whether it plays the important role in the household of Nature that its discoverers claim for it still has to be determined.

It has been reported that a large number of biological materials in which certain metabolic processes take place will emit this radiation; also that many materials in which these processes do not take place will emit the radiation as a secondary phenomenon when excited by primary radiation. Very often the most involved and complicated biological processes have been selected to work with. It would appear to the reviewers incautious at least to accept all this work, even if the fundamental phenomenon as such should exist. Where investigators of other problems have often spent years on careful preparation and extensive experiments, here the phenomenon has been solved after what often appears to be a preconceived pattern.

The micro-organisms which have been found to give up mitogenetic radiation are bacteria, yeast and fungi. Among the higher plants, the radiation is emitted by onion root tips and by ground up onion sole, from which "activator" or "mitotin" and an enzyme-like substance "mitotase" have been separated.[20, 24] It is also emitted by the leptom bundles in potatoes, by the mesophyll or leaf tissues of certain plants, and by beet tissues. It is here that the important fundamental work should have been done, but after the radiation was thought to have been established in these materials, they were completely neglected. Instead, the more thorough efforts were concentrated on animal materials, and again most often on phenomena of the utmost complexity. Sea urchin eggs in certain stages of development, and many other invertebrate materials are said to give up mitogenetic rays. Blood has been reported in many publications to be a good sender. This list of biological senders could be very much extended, but it is fully reported elsewhere.[20, 24]

A number of enzymatic reactions of fundamental importance, such as the action of rennet on milk, are reported to give mitogenetic rays (Fig. 2), and practically all chemical reactions are said to radiate. Wolff and Ras[61] report positive effects even with acid-base reactions and with dissolving electrolytes, but none with dissolving nonelectrolytes. An interesting discussion of the emission of short ultraviolet light (corresponding to mitogenetic rays) by electrochemical reactions is given by Braunstein and Potozky.[6] They report emission from concentration cells, reduction chains and electrolysis. The light always arises from the cathode, and the spectrum is usually independent of the electron donator but is characteristic of the electron acceptor.

Certain senders and detectors are usable only in visible light, but contradictory statements from different laboratories can be quoted. Thus Frank and Rodionov[14] wrote that the $K_2Cr_2O_7 + FeSO_4$ reaction will emit radiation only in the presence of visible light, while Barth[2] reported that this reaction will also emit mitogenetic rays in the dark.

VI. Advanced Work

The most important development, involving many contradictions, is that of the spectral analysis of mitogenetic radiation. A block of agar with a specially prepared yeast surface is placed beyond the prism of a medium size quartz spectrograph; that is, the photographic plate is replaced by a number of chambers containing yeast or other detectors, each covering a certain part of the spectrum. Or the spectrograph may be arranged as a monochromator, and behind the exit slit is placed a container with the detecting material. The detector can cover bands as narrow as 5A, and is exposed from 10 seconds to 20 minutes. After exposure, it is treated in the standard way and the increased growth is determined by the methods described above. Such work has shown that each chemical reaction emits its typical spectrum, although oxidation reactions have certain bands in common. The spectra of radiation emitted by biological materials contain bands of the chemical reactions taking place within them, and also bands which are not as yet identified, but which are possibly typical of the biological processes taking place. For example,[21] the radiation from nerves stimulated mechanically and from nerves stimulated

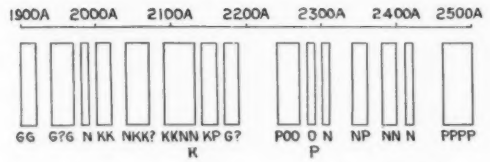


FIG. 6. Spectral analysis of mitogenetic radiation from rabbit blood, and identification of the bands. Redrawn from Golischewa.[16] *G*, Glycolitic spectrum; *N*, spectrum of peptid cleavages, so-called NH_3 spectrum, *K*, spectrum of creatin-phosphate cleavage; *P*, phosphatase spectrum; *O*, oxidation spectrum; *?*, lines of unknown origin.

electrically will differ in certain bands. It is thus possible to follow the chemical reactions taking place in biological processes without disturbing the living material, by simply identifying the bands emitted. Fig. 6 (redrawn from Golischewa[16]) illustrates the analysis of the radiation emitted by rabbit blood.

If this mitogenetic spectral analysis should be correct, it will open an entirely new approach to the study of biological problems, as well as provide a new tool for physical chemistry. But even if we should accept present detecting methods as satisfactory, and we seem still to be quite far from it, there are many objections from the physical point of view. It has been shown that sender cultures emit 10 to 1000 quanta/cm^2/sec., and the value seems rather to approach the lower limits. If now, one uses on the spectrograph a slit which permits the separation of 5A bands, and uses an exposure time of 10 seconds to 20 minutes, a simple calculation will show that only a few quanta would be available during the whole time of exposure. These few quanta (2 to 10) are supposed to stimulate a half cc of yeast suspension! This analysis of the radiation into 5A bands has been carried out down to 1900A . . . ![21]

There are clinical aspects to the problem which, if borne out by further investigation, may prove to be of the utmost importance. Mai[33] finds that the blood of healthy children will radiate, while that of rachitic children will not. The blood will resume its activity, however, when the rachitic condition has been removed. Gurwitsch[20] reports that the blood of animals will cease to radiate very shortly after the implantation of a cancer, although the cancer itself is supposed to be one of the most intense and constant sources of mitogenetic rays. Brainess[5] found that exhaustive labor temporarily destroys the power of the blood to radiate, although the statement is questioned by Hill.[23]

VII. CONCLUSION

We have attempted in the foregoing discussion to indicate to some extent the importance of the mitogenetic ray phenomenon in its physical, chemical, biological and clinical aspects. It is obvious from its present chaotic status, however, that more convincing, clear-cut evidence for the *existence* of the phenomenon must be established before it can be accepted as real. Even then, its full significance will not be realized until its detection is established as a routine laboratory procedure. It is the opinion of the reviewers that much could be accomplished toward clearing up the situation through the following program:

1. The methods which have been used in the detection of mitogenetic rays are such that they lend themselves to tests, and their limits of error can be definitely established. Therefore it should be possible for those who maintain a critical attitude but who have obtained opposite results, to work cooperatively on the fundamental problem until a conclusion has been reached on the following points: (a) Is short ultraviolet light emitted by biological materials and by chemical reactions? (b) Will artificial light of these intensities and wave-lengths stimulate growth in biological materials?

2. If and when the phenomenon is established, methods of detection should be developed which can be used as routine technique.

3. Of the 500 publications on the subject, obviously a large number are erroneous and contradictory. These should be considered carefully, and separated from those which have been submitted by the more responsible workers.

4. Future material for publication should be accepted only when sufficient details are given so that the experiments may be repeated by others.

Every new field attracts workers who believe that results may be obtained quickly and easily, and certainly this field has attracted more than its share of this type to both sides of the question. The mitogenetic ray problem is potentially too important to rest on the prejudices created by the irresponsible publications of these workers, and the reviewers feel that every effort should be exerted to put the problem on a substantial basis.

Bibliography

[1] R. Audubert and V. Doormaal, Comptes rendus **196**, 1883 (1933).

[2] H. Barth, Archiv Sci. Biol. Med. Exp. (Leningrad) **35**, 29 (1934).

[3] J. B. Bateman, Biol. Rev. **10**, 42 (1935).

[4] A. Becker, *Handbuch der Experimental Physik*, Vol. 23, (1928), pp. 1190–1205.

[5] S. Brainess, Arbeitsphysiol. **6**, 90 (1932).

[6] A. B. Braunstein and A. Potozky, Biochem. Zeits. **268**, 422 (1933).

[7] J. Chariton, G. Frank and N. Kannegiesser, Naturwiss. **18**, 411 (1930).

[8] W. W. Coblentz and H. R. Fulton, Bur. Standards Sci. Pap. **19**, 641 (1924).

[9] L. A. DuBridge, Phys. Rev. **37**, 392 (1931).

[10] J. R. Dunning, Rev. Sci. Inst. **5**, 388 (1934).

[11] J. Elster and H. Geitel, Physik. Zeits. **17**, 268 (1916).

[12] J. Eggert and W. Noddak, Naturwiss. **15**, 57 (1927).

[13] O. Ehrisman and W. Noethling, Zeits. f. Hyg. **113**, 597 (1932).

[14] G. Frank and S. F. Rodionov, Biochem. Zeits. **249**, 323 (1932).

[15] P. H. Geiger and W. Müller, Physik. Zeits. **29**, 839 (1928).

[16] K. P. Golischewa, Biochem. Zeits. **260**, 52 (1933).

[17] J. Gray and C. Ouellet, Proc. Roy. Soc. **B114**, 1 (1933).

[18] B. Gudden, *Lichtelektrische Erscheinungen* (J. Springer, 1928), p. 49.

[19] A. Gurwitsch, Archiv. mikrosk. Anat. u. Ent. mech. **100**, 11 (1923).

[20] A. Gurwitsch, *Die mitogenetische Strahlung* (J. Springer, 1932).

[21] A. and L. Gurwitsch, *L'Analyse Mitogenetique Spectrale* (Hermann et Cie, Paris, 1934).

[22] K. W. Hausser and K. H. Kreuchen, Zeits. f. tech. Physik **15**, 20 (1934).

[23] A. V. Hill, Nature **131**, 501 (1933).

[24] A. Hollaender, Chapter in *Biological Effects of Radiation*, Vol. II (B. M. Duggar, Editor), National Research Council (McGraw-Hill, 1935).

[25] A. Hollaender and W. D. Claus, Forthcoming publication.

[26] A. L. Hughes and L. A. DuBridge, *Photoelectric Phenomena* (McGraw-Hill, 1932), p. 171.

[27] A. L. Hughes and L. A. DuBridge, *Photoelectric Phenomena*, Chap. X.

[28] M. W. Karev and S. F. Rodionov, Zeits. f. Physik **92**, 615 (1934).

[29] K. H. Kreuchen and J. B. Bateman, Protoplasma **22**, 234 (1934).

[30] G. Locher, Phys. Rev. **42**, 525 (1932).

[31] W. Loos, Jahrbüch. f. wiss. Botanik **72**, 611 (1930).

[32] E. Lorenz, J. Gen. Physiol. **17**, 843 (1934).

[33] H. Mai, Archiv Sci. Biol. Med. Exp. (Leningrad) **35**, 237 (1934).

[34] S. Mardaschew and M. Mogliewsky, Biochem. Zeits. **265**, 28 (1933).

[35] L. Petri, Protoplasma **19**, 365 (1933).

[36] G. Protti, Radiobiologia **1**, 61 (1932).

[37] O. Rahn, Cold Spring Harbor Symp. Quant. Biol. **2**, 226 (1934).

[38] B. Rajewsky, Zeits. f. Physik **63**, 576 (1930). *Zehn Jahre Forschung auf dem physik-med. Grenzgebiet* (herausg. von F. Dessauer, Thieme, Leipzig, 1930), p. 244.

[39] B. Rajewsky, Ann. d. Physik **20**, 13 (1934).

[40] K. H. Reiss, Zeits. f. Physik **93**, 411 (1935).

[41] T. Reiter and D. Gabor, *Zellteilung und Strahlung* (J. Springer, 1928).

[42] W. Richards and G. Taylor, Biol. Bull. **63**, 113 (1932).

[43] B. Rossmann, Archiv f. Entwicklungsmech. **113**, 346 (1928).

[44] R. Ruyssen, Acta brev. neerl. Physiol. **3**, 141 (1933).

[45] O. H. A. Schmitt, Rev. Sci. Inst. **5**, 435 (1934).

[46] H. Schreiber and W. Friedrich, Biochem. Zeits. **227**, 386 (1930).

[47] H. Schreiber, Protoplasma **19**, 1 (1933).

[48] Note in Scientific American **152**, 43 (1935).

[49] F. Seyfert, Jahrbüch. f. wiss. Botanik **76**, 747 (1932).

[50] W. W. Siebert, Biochem. Zeits. **4**, 31 (1930); Strahlentherapie **40**, 780 (1931).

[51] W. W. Siebert and H. Seffert, Naturwiss. **21**, 193 (1933).

[52] W. Stempell, *Die unsichtbare Strahlung der Lebewesen* (G. Fischer, Jena, 1932).

[53] W. Stempell and G. Romberg, Protoplasma **13**, 28 (1931).

[54] R. Suhrmann, Zeits. f. Physik **33**, 63 (1925).

[55] G. W. Taylor and E. N. Harvey, Biol. Bull. **61**, 280 (1931).

[56] J. B. Tuthill and O. Rahn, Archiv Sci. Biol. Med. Exp. (Leningrad) **35**, 289 (1934).

[57] A. W. Uspensky, Zeits. f. Physik **40**, 456 (1926).

[58] L. K. Wolff and G. Ras, Ztrbl. f. Bakt. I. **123**, 258 (1931).

[59] L. K. Wolff and G. Ras, Ztrbl. f. Bakt. I. **128**, 305 (1933).

[60] L. K. Wolff and G. Ras, Archiv Scil. Biol. Med. Exp. (Leningrad) **35**, 51 (1934).

[61] L. K. Wolff and G. Ras, Biochem. Zeits. **250**, 305 (1932).

[62] R. W. G. Wyckoff, J. Gen. Physiol. **15**, 351 (1932).

[63] T. F. Young and W. C. Pierce, J. Opt. Soc. Am. **21**, 541 (1931).

An Experimental Determination of the Visual Thresholds at Low Values of Illumination

J. PRESTON CONNER AND ROBERT E. GANOUNG, *Massachusetts Institute of Technology*

(Received June 5, 1935)

INTRODUCTION

NUMEROUS data are available on visual acuity, the classical investigations of König having been supplemented by more recent work, perhaps the best of which is that of Lythgoe.[1] All such investigations are made with black objects against a white background, providing maximum possible contrast.

Unfortunately, the information obtained from such tests is of limited use because most practical problems deal with lesser contrasts. Cobb and Moss,[2] in their able research of 1928, pointed out this insufficiency of previous visual-acuity data and obtained results not only for one contrast but for a great range of contrasts. The work of Cobb and Moss, however, covers a range of luminosity* of only 1 to 100 lumens per square foot, and it was thought advisable to extend their results to lower values, including the whole range of luminosities used in street-lighting practice.

The present paper, therefore, gives the results of an investigation performed by the authors on the mutual relations among visual acuity, background luminosity and contrast.† The range of luminosities was from 0.0001 to 1.0 lumen per square foot, and the range of contrast was from 1.77 to 92.9 percent. The work was done in the Illuminating Engineering Laboratories of the Massachusetts Institute of Technology.

An important feature of this investigation was the determination of the relation between visual acuity and background luminosity for foveal vision alone and also for parafoveal vision alone. The use of a fixation point allowed the parafoveal curves for visual acuity *vs.* luminosity to be extended from 0.0001 to 1.0 lumen per square foot, and these curves were found to be practically straight lines of small slope. Similarly, the curves for foveal vision were found to remain straight and to descend sharply to zero as the luminosity was reduced. It is believed that the present work constitutes the first such separation to be made, and it is hoped that this information will be of value in the further development of theories of vison.

APPARATUS

The illumination was produced by tungsten-filament lamps operating at a color temperature of approximately 2670 degrees K. A series of lamps of various sizes was used, operated at such values of current as to give essentially the same color temperature in all cases. The light was diffused by a sheet of opal glass which was held in a special luminaire shielded from view. The luminaire was so placed that specular reflections from test objects and background were entirely eliminated. All reflecting surfaces were also made as nearly perfectly diffusing as possible. In the tests with red light, the filament temperatures varied considerably. However, this fact is unimportant since the radiation was restricted to a narrow band, a filter being used whose transmission factor was less than 0.50 percent at a wave-length of 0.60 micron.

The standard test object or broken circle was used throughout the investigation. A large number of brass disks were covered with a white water paint which produced a very good matt background with a reflection factor of 87 percent. On these disks were glued the test objects which were made from papers having the desired reflection factors to give the several contrasts. A

[1] R. J. Lythgoe, *The Measurement of Visual Acuity*, Medical Research Council, Special Report No. 173, 1932.

[2] P. W. Cobb, and F. K. Moss, *Four Fundamental Factors in Vision*, Trans. I. E. S. **23**, 496 (1928); *The Four Variables of the Visual Threshold*, J. Frank. Inst. **205**, 831 (1928).

* Luminosity is defined as the total luminous flux emitted per unit area of source. It is expressed in lumens per square foot throughout the paper. For any surface, luminosity (L) is equal to illumination (E) times reflection factor (ρ), or $L=\rho E$. For perfectly diffusing surfaces, luminosity (lumens/ft.^{2}) is equal to brightness in so-called "foot-lamberts."

† Contrast is defined as $(L_B-L_0)/L_B$, where L_B=luminosity of the background and L_0 is the luminosity of the object. Generally this is expressed in percent or $(L_B-L_0)/L_B\times 100$. It is considered adequate only when $L_B>L_0$, which was the case throughout this investigation.

complete set (fourteen different sizes) of test objects was made for each contrast. The test objects were placed against a white surround which covered the entire end of the room (Fig. 1). The luminosity of the surround was that which resulted from using only the one luminaire required for the illumination of the background and test object. The luminosity, however, was fairly uniform, and according to Lythgoe[1] the effect of small variations is negligible.

Experimental Procedure

Background luminosity, contrast and visual acuity were the only variables in this work. The exposure time was of the order of three seconds or more and hence was essentially equal to infinite exposure time as far as the results are concerned.[3] Values of luminosity are in the ordinary "photopic lumens" per square foot throughout, care being taken to avoid "scotopic lumens" even at the lowest levels of illumination. Thus all the curves could be plotted with energy units as abscissae without altering their shapes. Binocular vision was employed throughout the investigation, and the natural pupil was used with no correction.

The readings were taken in periods of approximately three hours, with frequent rest periods. Before each run and after a change in background luminosity, time was allowed for adaptation to the particular luminosity being used. This adaptation period was from 30 to 40 minutes, depending on the luminosity and somewhat on the condition of the observer's eyes.

The gap in the test object was placed in one of eight possible positions. For each exposure of an object the observer was required to judge the position of the gap, even though the judgment was apparently a guess. The order of placing the gap in the various positions was haphazard. Furthermore, the various contrasts were intermixed.

The determination of the relation between visual acuity and background luminosity for parafoveal vision required control of fixation.

Fig. 1. The observer and room taken under actual operating conditions with a background luminosity of one lumen per sq. ft. A test object is shown in position and the spare objects are held by the rack on the left.

Usual incandescent-lamp light was used, but a small red fixation lamp was placed 5 degrees above the center of the test object. The observer was required to keep his eyes focused on this point while the test object was exposed.

In analyzing the data to establish the visual threshold, the condition of 4.5 correct answers out of 8.0 trials was used as the criterion. This procedure was used by Lythgoe,[1] who discusses the matter in considerable detail. For each luminosity and contrast, a curve was plotted with the angular size of the test object in minutes as abscissa and with the number of correct answers out of 8.0 trials as ordinate. From these curves the angular size in minutes corresponding to 4.5 correct answers was obtained from each observer. The reciprocal of this angular size was taken as the visual acuity. The curve shown in Fig. 2 illustrates the method used. It is interesting to note that the curves of angular size of the test object against the number of correct answers are very nearly straight lines in the central region, but are curved at the ends so that they are somewhat *S*-shaped. The data as a whole show that, for constant contrast, the slope of the curves decreases as the luminosity is decreased. These two characteristics are essentially the same as found by Lythgoe.[1] An additional tendency noted in this investigation is that at constant luminosity the slope of the curves decreases somewhat with a decrease in contrast.

[3] According to Ferree and Rand, the effect of increases in exposure time on visual acuity is almost negligible with exposure times greater than 0.10 second for a background luminosity 5 lumens per square foot. C. E. Ferree and Gertrude Rand, Trans. I. E. S. **23**, 521 (1928).

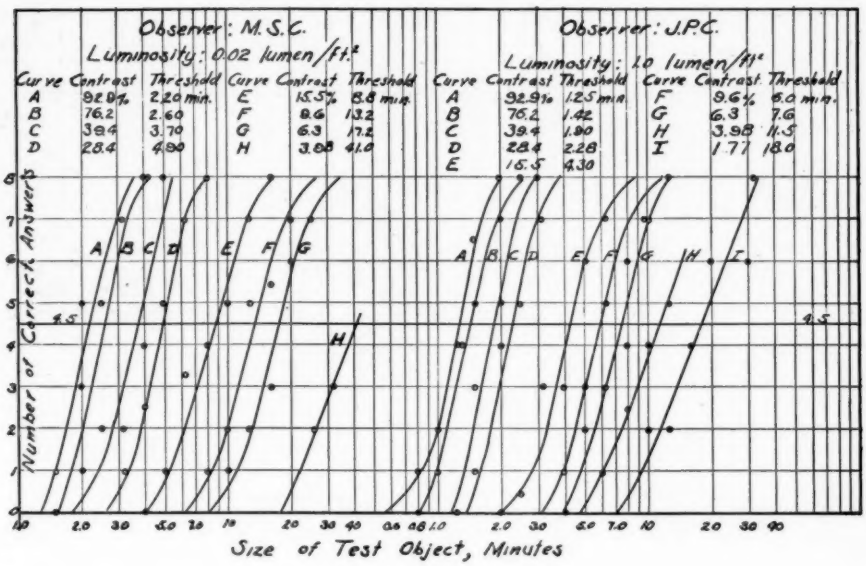


FIG. 2.

Seven normal observers were used in determining the mutual relations among visual acuity, background luminosity and contrast. However, only two of these observers were used in determining the relation between visual acuity and background luminosity for foveal and parafoveal vision. Approximately 29,000 readings were secured in this investigation. In averaging the results to obtain the final values the geometric average was used.

RESULTS

The results of this investigation for incandescent-lamp light show the mutual relations among luminosity of the background, contrast and visual acuity. Table I gives the complete average data for these mutual relations while Table II gives the same results but for the two observers and four contrasts used in determining the relations for foveal and parafoveal vision. Table III shows the average results for two observers using red light. Table IV gives the average results for the offset-fixation method.

The curves of visual acuity against background luminosity for constant contrast, Fig. 3, are straight lines in the upper portion, that is, at luminosities above approximately 0.017 lumen

TABLE I. *Geometric average values of visual acuity.* Seven observers; incandescent lamps.

Percent Contrast	Luminosity of background, lumens per square foot							
	0.00013	0.000983	0.00993	0.020	0.044	0.100	0.312	1.00
92.9	0.055	0.114	0.335	0.459	0.638	0.741	0.850	0.965
76.2	0.044	0.089	0.268	0.404	0.508	0.657	0.715	0.863
39.4	0.030	0.055	0.194	0.267	0.370	0.443	0.538	0.632
28.4	0.023	0.041	0.131	0.198	0.291	0.359	0.464	0.584
15.5	—	0.025	0.071	0.105	0.160	0.198	0.258	0.335
9.6	—	—	0.040	0.064	0.114	0.125	0.158	0.203
6.3	—	—	0.034	0.054	0.087	0.119	0.141	0.186
3.98	—	—	—	0.036	0.039	0.048	0.060	0.085
1.77	—	—	—	—	0.028	0.033	0.046	0.071

TABLE II. *Geometric average values of visual acuity.* Two observers; incandescent lamps.

Percent Contrast	Luminosity of background, lumens per square foot							
	0.00013	0.000983	0.00993	0.020	0.044	0.100	0.312	1.00
92.9	0.040	0.090	0.396	0.515	0.609	0.710	0.960	0.930
76.2	0.034	0.073	0.350	0.427	0.501	0.663	0.849	0.865
39.4	0.028	0.043	0.188	0.309	0.340	0.417	0.628	0.689
28.4	0.014	0.032	0.139	0.220	0.268	0.348	0.529	0.633

TABLE III. *Geometric average values of visual acuity.* Two observers; red light.

Percent Contrast	Luminosity of background, lumens per square foot						
	0.00076	0.0042	0.0064	0.010	0.018	0.031	0.076
92.8	—	0.075	0.126	0.198	0.335	0.421	0.520
73.8	—	0.043	0.097	0.168	0.286	0.334	0.480
36.5	—	0.021	0.046	0.102	0.159	0.197	0.322
26.6	—	—	0.030	0.063	0.120	0.145	0.205

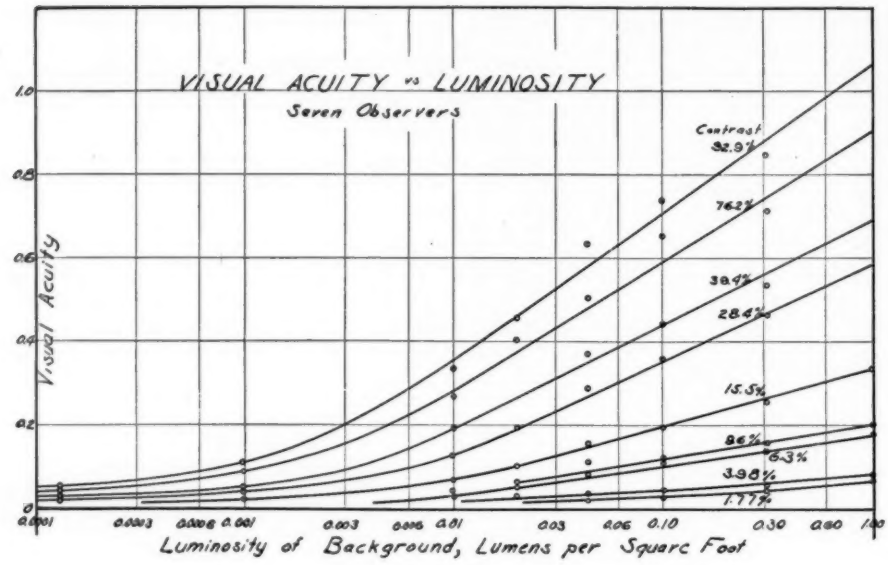


FIG. 3.

TABLE IV. *Geometric average values of visual acuity.* Two observers; offset fixation [5° from fovea].

Percent Contrast	Luminosity of background, lumens per square foot							
	0.00013	0.000983	0.00993	0.020	0.044	0.100	0.312	1.00
92.9	0.058	0.10	0.146	0.162	0.189	0.227	0.252	0.289
76.2	0.046	0.089	0.119	0.131	0.162	0.187	0.225	0.234
39.4	0.028	0.053	0.101	0.105	0.113	0.113	0.138	0.154
28.4	—	0.040	0.070	0.093	0.083	0.112	0.105	0.121

per square foot. In the region of lower luminosity the curves are no longer straight lines but become curved. The straight-line portion at the high luminosities represents the predominating effect of cone vision. In the region from approximately 0.002 to 0.02 lumen per square foot both the rods and cones are operative. Below 0.002 lumen per square foot, vision is a function of the rods alone and the curves again approach straight lines.

For the purpose of comparing the results of this investigation with those of well-known works of the past, Fig. 4 has been prepared. It shows the results obtained by Lythgoe,[1] Cobb and Moss,[2] and König[4, 5] in similar experiments. The only curve shown in Fig. 4 from our investigation is that for 92.9 percent contrast. Lythgoe's curve is for an exposure time of 1.6 seconds and a contrast of 98.1 percent. The curve determined by the data of Cobb and Moss is shown dashed because it is merely a straight-line extension of the curve which they obtained and which was very nearly a straight line in the interval from 1.0 to 100 lumens per square foot. This curve was obtained for an exposure time of 0.170 second and a contrast of 100 percent. König's curve is for the so-called contrast of "black object on a white background," and the exposure time is indefinite. The differences evidenced in Fig. 4 are presumably due largely to variations in observers; through difference in exposure time, contrast and the surround may have a slight effect.

The curves of Fig. 5 showing visual acuity as a function of percent contrast for particular luminosities are merely cross-plots of the curves of Fig. 3.

RESULTS USING RED LIGHT

The results for red light, Fig. 6, show the variation of visual acuity with luminosity for four particular contrasts. The purpose of using red light was to determine the relation of visual

[4] A. König, *Die Abhängigkeit der Sehschärfe von der Beleuchtungsintensität*, Gesammelte Abhandlungen zur Physiologischen Optik (1903).

[5] S. Hecht, *The Relation between Visual Acuity and Illumination*, J. Gen. Physiol. **2**, 255 (1927–29).

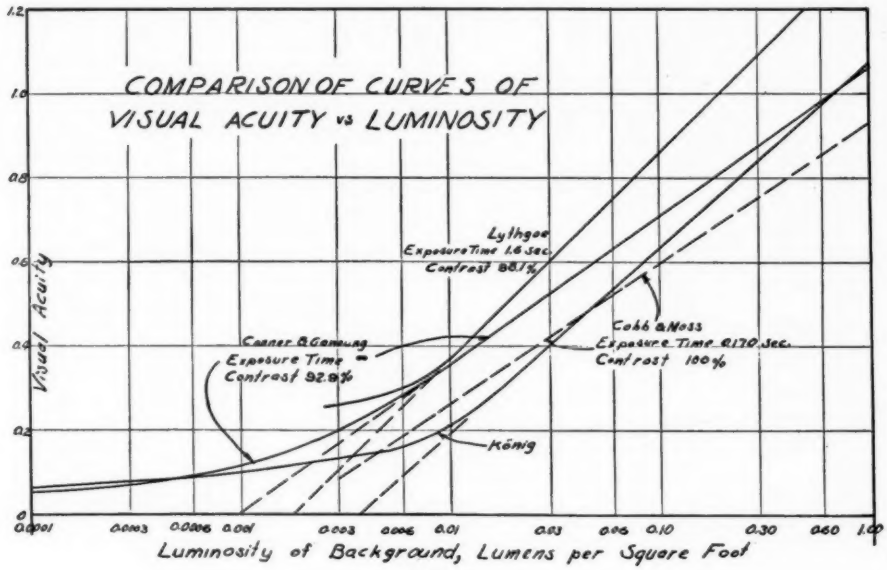


FIG. 4.

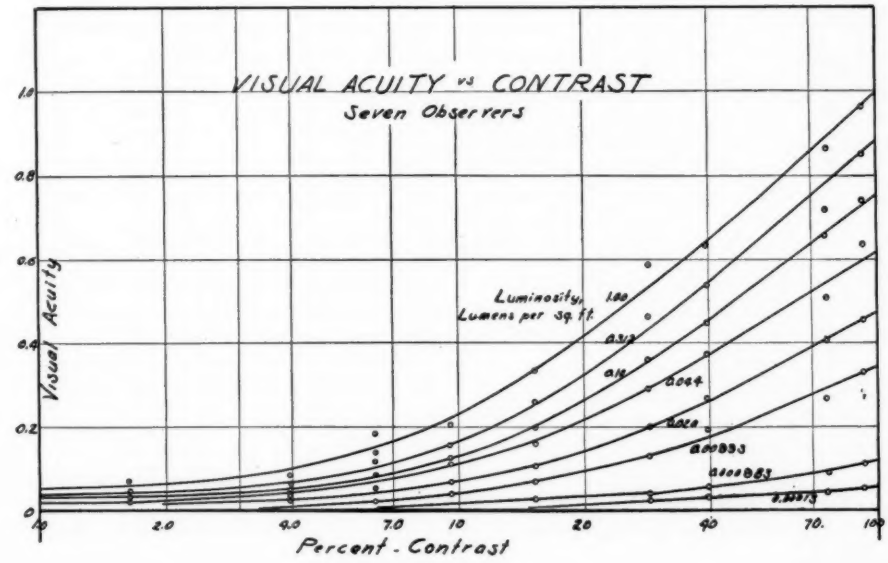


FIG. 5.

acuity to background luminosity for foveal vision only. The duplicity theory of vision assumes the rods and cones to be two elements with distinct functions in vision, the visibility curve[6] for the rods being similar to that for the cones but shifted toward the blue end of the spectrum. Thus by using red light with as little energy as

[6] S. Hecht and R. E. Williams, *The Visibility of Monochromatic Radiation and the Absorption Spectrum of Visual Purple*, J. Gen. Physiol. **5**, 1 (1922).

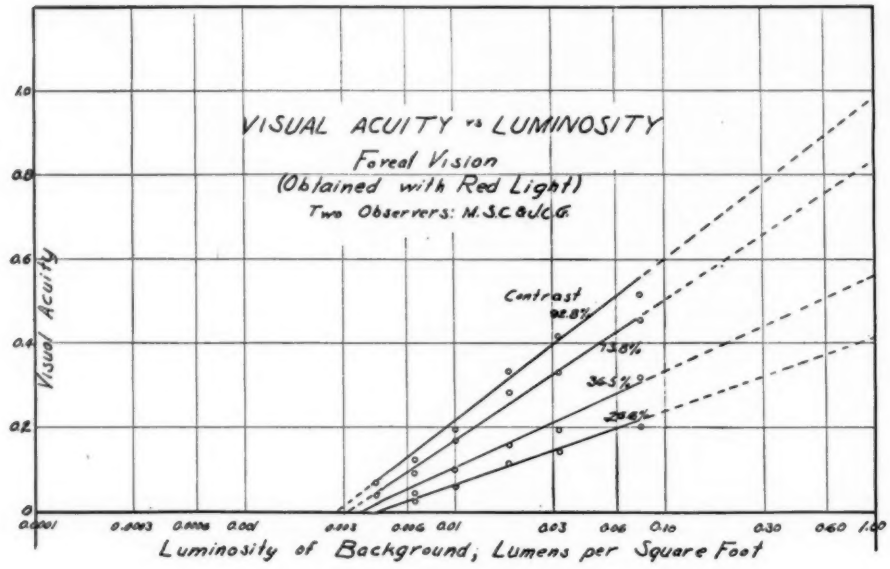


FIG. 6.

possible below a wave-length of 0.62 micron, the cones are operative while the rods are practically unexcited. The data of Fig. 6 were taken for only two observers. In order to compare the results with those obtained in incandescent-lamp light (filament operating at, 2670°K) the results for the same two observers for white light are tabulated in Table II and plotted in Fig. 7.

The curves for red light are straight lines throughout their length. It is definitely known that the curves continue to the axis of zero visual acuity because at a luminosity of 0.00076 lumen per square foot the visual acuity was so small that it could not even be estimated with the test objects available. Because of experimental difficulties no points were obtained between 0.00076 and 0.0042 lumen per square foot, but it is reasonable to expect the curves to continue as straight lines to the axis. The curve for 26.6 percent contrast is very close to the axis of zero visual acuity at 0.0042 lumen per square foot, as may be seen from the data. The curves obtained with red light are very nearly parallel to the straight-line (foveal) portions of the curves for white light. This fact indicates that the visual acuity of the cones bears a strictly linear relation to the background luminosity from their absolute threshold of approximately 0.002 lumen per square foot to the highest values used in our investigation.

RESULTS FOR PARAFOVEAL VISION

In this case as in the previous ones, only four contrasts were used and the results obtained for two observers. The geometric averages are shown in Table IV and also in the curves of Fig. 8. The use of the offset-fixation method permitted the separate determination of the parafoveal function of visual acuity with respect to the background luminosity. The eyes of the observer were kept focused on the small red fixation lamp so that its image was in the fovea. The fact that the light was red exerted an influence to prohibit peripheral fixation. The observers in all cases said that it was easy to keep the eyes focused on the fixation light when the test objects were exposed to view. The curves for the parafovea, Fig. 8, are straight lines to the region of lower luminosities, that is, up to 0.02 lumen per square foot. The curves for the two lower contrasts are straight lines throughout their entire length. The slight upward bend in the curves for the two higher contrasts is very probably due to a small number of cones which were activated by the peripheral image of

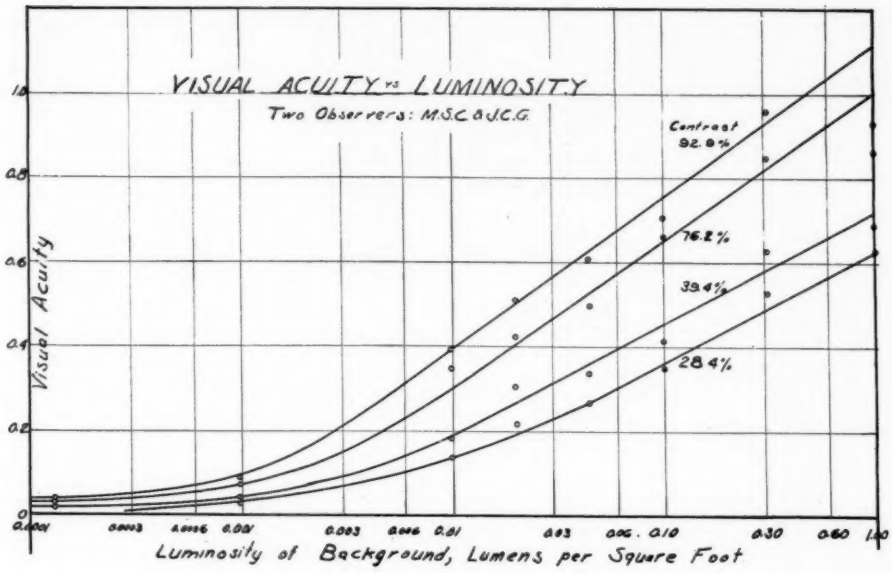


FIG. 7.

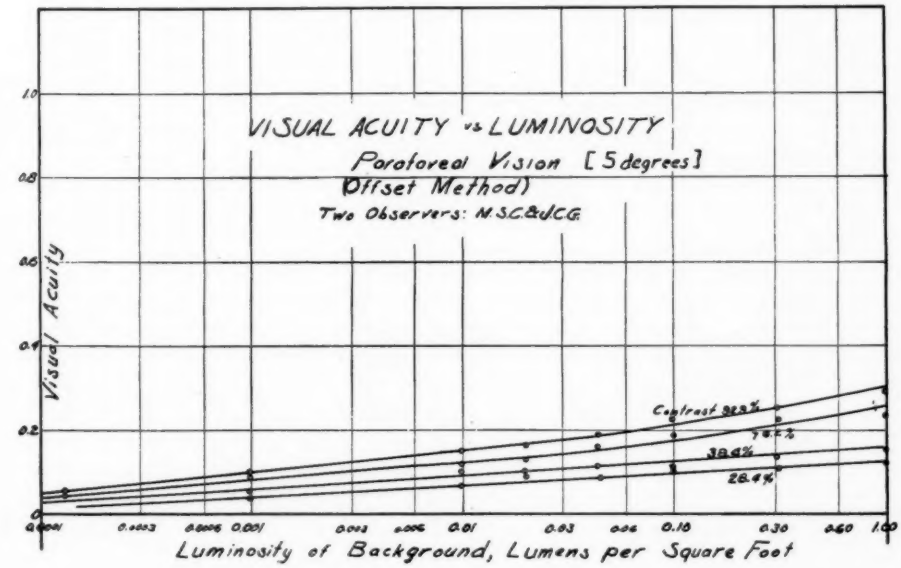

VISUAL ACUITY vs LUMINOSITY
Parafoveal Vision [5 degrees]
(Offset Method)
Two Observers: M.S.C. & J.C.G.
Visual Acuity
Luminosity of Background, Lumens per Square Foot

FIG. 8.

the test object. The data indicate that the visual acuity of the rods also bears a linear relation to the logarithm of the background luminosity, though the slope of the curves is much less than for foveal vision.

According to the duplicity theory, the eyes of a totally color-blind person are devoid of cones. The visual acuity curves in such cases should be determined by the activity of the rods alone. Fig. 9 shows the relation between visual acuity

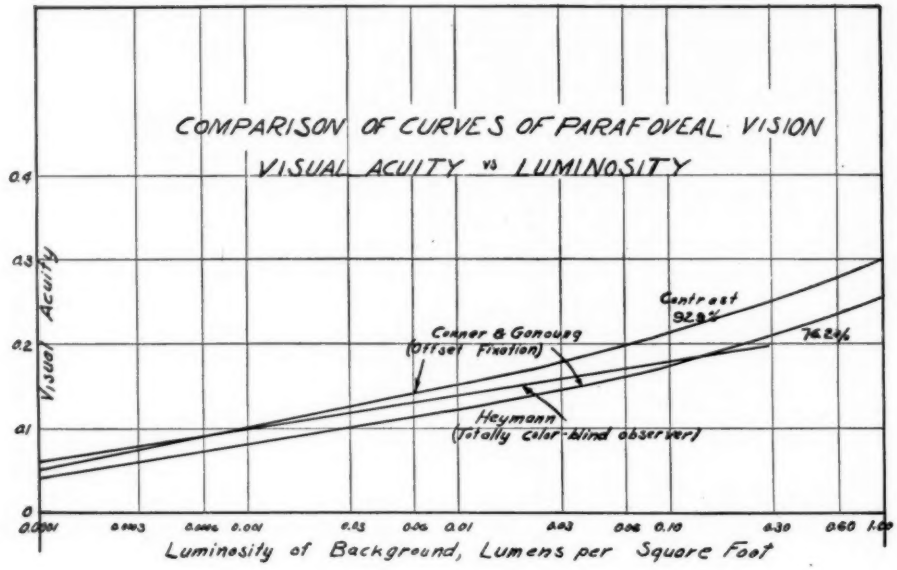


FIG. 9.

and background luminosity obtained by König[4] for a totally color-blind individual. In this same figure are shown the two curves obtained in this investigation for the two higher contrasts. The agreement between the curves is striking.

CONCLUSIONS

The conclusions may well be divided into three parts: (1) normal vision, (2) foveal vision and (3) parafoveal vision.

(1) For normal vision, curves have been obtained which indicate the relations between contrast and visual acuity for luminosities varying from 0.0001 to 1.0 lumen per square foot.

(2) Visual-acuity curves were determined for four contrasts, and with background luminosities varying from 0.003 to 1.0 lumen per square foot, the effect of rod vision being minimized by using red light.

(3) For these same four contrasts visual acuity curves were obtained for vision five degrees from the fovea. The range of luminosities for this parafoveal vision was from 0.0001 to 1.0 lumen per square foot.

The visual-acuity curves for normal vision in light from incandescent lamps give specific information applying to the low values of illuminations encountered in highway and outdoor lighting. The results show that foveal and parafoveal vision are governed by separate linear functions. These facts suggest that separate consideration of these two linear functions of visual acuity might result in simpler and more exact methods of treating illumination problems.

The authors wish to express their sincere appreciation to Professor Parry Moon of the Massachusetts Institute of Technology for his valuable advice.

On the Fundamental Optical Invariant, the Optical Tetrality Principle, and on the New Development of Gaussian Optics Based on This Law*

M. HERZBERGER, *Eastman Kodak Company, Rochester, N. Y.*

(Received June 27, 1935)

THE author has developed in his book *Strahlenoptik*, the laws of geometrical optics, using the principle of Fermat. This procedure presupposes some acquaintance with the methods of the variation calculus. In the present paper it will be shown how the optical laws may be derived from the law of refraction (Snellius-Descartes). The author shows in the first part of this paper that there is for any two-dimensional manifold of rays a certain differential invariant (first published in *Strahlenoptik*, p. 13). His intention is to show that this fundamental law includes all other known optical laws as special cases, and it may therefore be used as a starting point in developing geometrical optics systematically.

The value of this method will be shown here by its application to the problems of Gaussian optics. We shall develop systematically all the theorems of Gaussian optics according to the methods sketched by the author[1] whereby new connections are found between these laws. This method of investigation will be applied in a second paper to the investigation of nonsymmetrical systems dealt with by A. Gullstrand; a third paper will deal with the laws obtained by integrating the differential invariant.

At the end of the paper the author will group all the objections to the method of Gaussian optics as usually treated. The letters used in this paper and the notations are in accord with the German standard report DIN1335, *Bezeichnungen in der technischen Optik* (Notations in Technical Optics).

The author is very much obliged to Professor Kingslake and Dr. Webb for looking after the linguistic part of this paper.

I. The Law of Refraction and the Derivation of the Differential Invariant

A light ray may be defined by one of its points, $\mathbf{a}$, and the unit vector, $\mathbf{s}$, in its positive direction. The corresponding image ray, produced by an arbitrary optical system, might in the same way be specified by the vector to one of its points, $\mathbf{a}'$, and the unit vector, $\mathbf{s}'$, in its direction. The origin of the coordinates in object and image spaces may be chosen arbitrarily.

Let n be the refractive index of the object space, n', of the image space. Then we search for a function of n, $\mathbf{a}$ and $\mathbf{s}$, which is equal to the same function, written in n', $\mathbf{a}'$ and $\mathbf{s}'$.

To find this law, we analyze the refraction at a single surface.

Let $\mathbf{o}$ (Fig. 1) be the unit vector normal to the refracting surface at the point of incidence of our ray, $\mathbf{s}$, $\mathbf{s}'$, the unit vectors of the ray before and after refraction. We can then write the Snellius-Descartes law of refraction in the form

$$n\mathbf{s}\times\mathbf{o}=n'\mathbf{s}'\times\mathbf{o}'. \tag{1}$$

Here, the sign $\times$ means the vectorial multiplication of the two vectors. We can write instead of (1)

$$(n'\mathbf{s}'-n\mathbf{s})\times\mathbf{o}=0,$$

which means that the two vectors have the same direction. We can therefore write

$$(n'\mathbf{s}'-n\mathbf{s})=\Gamma\mathbf{o}. \tag{2}$$

The proportionality factor is easily calculated. By multiplying (2) scalarly by $\mathbf{o}$, we find

$$\Gamma=n'\cos\epsilon'-n\cos\epsilon,$$

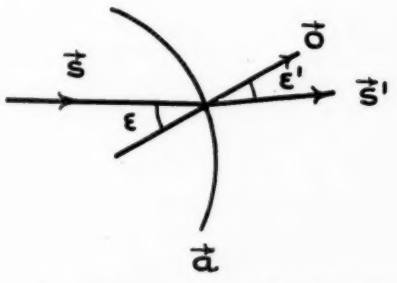


FIG. 1. Refraction at a single surface $\vec{a}$ (U_1V) $\vec{o}$ unit vector normal to the surface $\vec{s}$, $\vec{s}'$ unit vectors along the ray before and after refraction.

* Communication No. 555 from the Kodak Research Laboratories.

[1] Herzberger, Physica **2**, 239 (1935).

FIG. 2. Shifting of a line element along two neighboring rays from A to B.

ϵ, ϵ' being the angles of the normal $\mathbf{o}$ with the incident and refracted rays, respectively. We do not need, however, to consider this in the present paper since we do not use the actual value of Γ.

Now we shall investigate a two-dimensional manifold of incident rays, for which $\mathbf{a}$ and $\mathbf{s}$ depend on two parameters, e.g., u and v, i.e., in general, $\mathbf{a}$ is a surface such that through each point of the surface passes one ray, defined by $\mathbf{a}$ and its direction given by the unit vector $\mathbf{s}$. We denote $\partial\mathbf{a}/\partial u$ by the abbreviation $\mathbf{a}_u$. Then by the definition of $\mathbf{o}$,

$$\mathbf{o}\cdot\mathbf{a}_u=\mathbf{o}\cdot\mathbf{a}_v=0. \tag{3}$$

Differentiating these two equations with respect to v, and u, we find that

$$\mathbf{o}_v\mathbf{a}_u=\mathbf{o}_u\mathbf{a}_v=-\mathbf{o}\mathbf{a}_{uv}. \tag{4}$$

If we differentiate Eq. (20) with respect to u and v, and consider Eq. (3), after multiplying by $\mathbf{a}_v$ and $\mathbf{a}_u$, respectively, we get

$$n'\mathbf{s}_u'\mathbf{a}_v'-n\mathbf{s}_u\mathbf{a}_v=\mathbf{o}_u\mathbf{a}_v,$$
$$n'\mathbf{s}_v'\mathbf{a}_u'-n\mathbf{s}_v\mathbf{a}_u=\mathbf{o}_v\mathbf{a}_u.$$

Then, by subtraction, and including Eq. (4), we find that

$$n'(\mathbf{s}_u'\mathbf{a}_v'-\mathbf{s}_v'\mathbf{a}_u')=n(\mathbf{s}_u\mathbf{a}_v-\mathbf{s}_v\mathbf{a}_u)=I. \tag{5}$$

We see that the magnitude of I does not vary upon refraction.

We shall now show that I is an invariant, which does not depend on the choice of the intersecting surface $\mathbf{a}$ made above. If $\mathbf{b}$ is another surface intersecting our rays, we can write the equation

$$\mathbf{b}=\mathbf{a}+\lambda\mathbf{s}, \tag{6}$$

in which ∂ also can vary from ray to ray, and must therefore be regarded as a function of u and v. By differentiating we get

$$\mathbf{b}_u=\mathbf{a}_u+\lambda_u\mathbf{s}+\lambda\mathbf{s}_u,$$
$$\mathbf{b}_v=\mathbf{a}_v+\lambda_v\mathbf{s}+\lambda\mathbf{s}_v.$$

Multiplication by $n\mathbf{s}_v\times n\mathbf{s}_u$, respectively, gives, on subtraction,

$$n(\mathbf{b}_u\mathbf{s}_v-\mathbf{b}_v\mathbf{s}_u)=n(\mathbf{a}_u\mathbf{s}_v-\mathbf{a}_v\mathbf{s}_u). \tag{7}$$

If we take the theorems together, we can demonstrate the following *theorem of the fundamental invariant:* Given a two-dimensional manifold of rays, every ray by one of its points, $\mathbf{a}$, and the unit vector in its direction, $\mathbf{s}$, and in the image space by $\mathbf{a}'$ and $\mathbf{s}'$; then the expression

$$I=n(\mathbf{s}_u\mathbf{a}_v-\mathbf{s}_v\mathbf{a}_u)=n'(\mathbf{s}_u'\mathbf{a}_v'-\mathbf{s}_v'\mathbf{a}_u') \tag{8}$$

is an invariant, independently of how the manifold of rays is intersected. I is a differential invariant and each invariant for the given manifold can be derived from it.

We can state our law in a more convenient form if we consider instead of the unit vectors $\mathbf{s}$ and $\mathbf{s}'$ the vectors $\mathbf{n}=n\mathbf{s}$ and $\mathbf{n}'=n'\mathbf{s}'$, the lengths of which are equal to the refractive indices of the respective media. Then our law takes the following form:

$$\mathbf{n}_u\mathbf{a}_v-\mathbf{n}_v\mathbf{a}_u=\mathbf{n}_u'\mathbf{a}_v'-\mathbf{n}_v'\mathbf{a}_u'. \tag{8a}$$

This form of the general law suggests a kind of reciprocity law. Let us replace in (8a)

(1) $\mathbf{n}$ by $\mathbf{a}$, $\mathbf{a}$ by $-\mathbf{n}$ or (2) $\mathbf{n}'$ by $\mathbf{a}'$, $\mathbf{a}'$ by $-\mathbf{n}'$,

(3) $\mathbf{n}$ by $\mathbf{a}$, $\mathbf{n}'$ by $\mathbf{a}'$, $\mathbf{a}$ by $-\mathbf{n}$, $\mathbf{a}'$ by $-\mathbf{n}'$.

Eq. (8a) does not vary if we make any of the three substitutions suggested above.

We may therefore conclude: From all optical formulas developed by means of (8) we derive three new correct formulas if we express in it all magnitudes with the help of $\mathbf{a}$ and $\mathbf{n}$; we obtain three new correct equations if we make one of these three substitutions. This principle will be employed often in the following discussions.

It is convenient to write the equation somewhat differently. We shall use differentials instead of differential quotients.

Given three rays, infinitely near together, namely, one principal ray through $\mathbf{a}$ and $\mathbf{s}$, and the neighboring rays through $\mathbf{a}+d_1\mathbf{a}$, $\mathbf{s}+d_1\mathbf{a}$, and $\mathbf{a}+d_2\mathbf{a}$, $\mathbf{s}+d_2\mathbf{s}$.

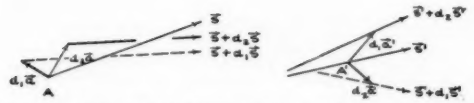


FIG. 3. Three neighboring rays in arbitrary situation in object and image space.

Eq. (5) gives for these three rays

$$n'(d_1\mathbf{s}'d_2\mathbf{a}'-d_2\mathbf{s}'d_1\mathbf{a}')=n(d_1\mathbf{s}d_2\mathbf{a}-d_2\mathbf{s}d_1\mathbf{a}), \quad (9)$$

which is valid for all rays in the neighborhood of the principal ray as long as they are so near that the Taylor development may be broken off after the linear element.

We now suggest a third form. If we know the image ray of one neighboring ray ($d\mathbf{s}_1$, $d\mathbf{a}_1$; $d\mathbf{s}_1'$, $d\mathbf{a}_1'$) then Eq. (9) shows that, for an arbitrary neighboring ray, the equation

$$n'(d\mathbf{s}'d\mathbf{a}_1'-d\mathbf{s}_1'd\mathbf{a}')=n(d\mathbf{s}d\mathbf{a}_1-d\mathbf{s}_1d\mathbf{a}) \quad (10)$$

is correct. If the image of *one* neighboring ray is known, then (1) tells something about every neighboring ray. We shall use the fundamental equation most frequently.

II. GAUSSIAN OPTICS

(a) The fundamental law

Given an optical system with axial symmetry we will study the rays in the neighborhood of the axis.

Because the system is symmetrical we only need to investigate the rays in a meridian plane. The image of a skew ray can be obtained by projecting it on two meridian planes. The images of the projections then are the projections of the image ray. A demonstration of this theorem, first stated by Lippich, involves no optical laws, only the fact that the system has rotational symmetry.

A meridian ray may be indicated by the height* $\bar{h}$ (Fig. 4) by which it intersects a line orthogonal to the optical axis, and by its angle σ with the axis; analogously by h' and σ' in the image space.

If we want to apply the theorem we must assume the hypothesis that h and σ are small quantities. σ will be given in radian measure, and is so small that we may substitute the angle for its sine.

* *Note:* We write conjugate magnitudes with the same symbol distinguishing the image magnitude by a prime. If we write not conjugated magnitudes, with the same letter, we write it with a bar, e.g., $\bar{h}$, h'; $\bar{f}$, f', $\bar{g}$, g'.

FIG. 4. The determining elements of a meridian ray in Gaussian optics. $\bar{O}$, O' origin of coordinates.

FIG. 5. Two rays through the same point of the axis.

The following laws and the whole field of Gaussian optics hold if and only if the hypothesis mentioned above is allowed.

In this case **da** is orthogonal to the axis and has the length h. **ds** has the same direction and the length σ. The fundamental theorem takes the form

$$n'(h_1'\,\sigma'-h'\sigma_1')=n(\bar{h}_1\,\sigma-\bar{h}\sigma_1). \quad (2,1)$$

Our tetrality principle allows in each valid equation the substitutions

(1) $n\sigma\rightarrow\bar{h}$, $\quad\bar{h}\rightarrow-n\sigma$,

(2) $n'\sigma'\rightarrow h'$, $\quad h'\rightarrow-n'\sigma'$,

(3) $n\sigma\rightarrow\bar{h}$, $\quad n'\sigma'\rightarrow h'$, $\quad\bar{h}\rightarrow-n\sigma$, $\quad h'\rightarrow-n'\sigma'$.

(b) The points on the axis have a sharp image

The invariant vanishes if

$$\sigma_1 : h_1=\sigma : h. \quad (2,2)$$

Then, we have

$$\sigma_1' : h_1'=\sigma' : h'. \quad (2,3)$$

The geometrical meaning of (2, 2) is that we are considering all rays that meet the axis at the same point as our first ray. Then (2, 3) shows that the rays converge also in the image space at the same point.

Theorem I. The rays emerging from a point on the axis converge in the image space at a point on the axis.

These points on the axis have, within the limit of Gaussian optics, a sharp image.

This theorem is also correct for the focal point if we consider a bundle of parallel rays as converging to an infinite point.

As regards the infinite point in the object space, two cases are possible. First, the image point is a finite point, called the *focal point* of the image space. Second, it is infinite, in which case

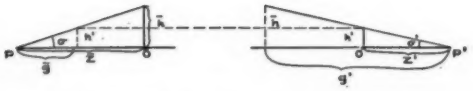


FIG. 6. Definition of the apparent distance of the object and image origin $\bar{g}$, respectively, g'.

we call the system *afocal*. All telescope systems are *afocal*. Hence, we sometimes call these systems *telescopic systems*. The *focal point* of the object space is the point whose image is the infinite point of the image space.

(c) Some expressions of optical value

A ray is indicated by h, $n\sigma$, h', $n'\sigma'$. The ratio of every two of these expressions has a significant optical meaning and we shall denote each meaning by a symbol.

First, $h/\sigma=\bar{z}$, $h'/\sigma'=z'$ gives the distance of the origin from the point where the ray meets the optical axis.

Then

$$h'/\bar{h}=\delta', \quad \sigma'/\sigma=\gamma'. \tag{2, 4}$$

If we place our eye at the point P, then the rays project our image coordinate plane on the object coordinate plane. This projection gives a similar one to one point correspondence and its magnification may be called the projection coefficient, according to A. Gullstrand. We denote the *projection coefficient* by the letter δ'.

γ' is the *angle coefficient*, usually designated as angular magnification. Further we put

$$h'/\sigma=\bar{g}, \quad \bar{h}/\sigma'=g'. \tag{2, 5}$$

The optical meaning of these quantities is as follows:

According to R. Smith $\bar{g}$ is the *apparent distance* of the coordinate plane in the image seen from P.

If we place our eye at the point P and view the image plane we see all objects in this plane through our optical system under the same angle as if the object were at a distance $\bar{g}$ without any optical system.

Similarly, we define the meaning of g'. It is the apparent distance of the object coordinate plane seen from P' with the aid of our optical system.

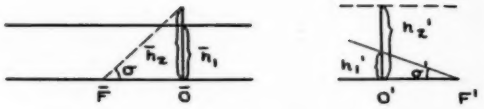


FIG. 7. Two rays, one entering parallel to the axis, the other sorting parallel to the axis. Figure to demonstrate the equation given by Gauss between the focal lengths.

From the definitions of our expressions:

$$\bar{g}=\gamma' z', \quad z'=g'\delta', \quad g'\gamma'=\bar{z}, \quad \bar{z}=\bar{g}/\delta', \tag{2, 6}$$

and

$$\bar{g}=\gamma'\delta' g'.$$

The above section dealt only with definitions; no single optical law was involved.

(d) Some relations between the fundamental expressions introduced above

We remember that our origins in object and image space have arbitrary positions; they are therefore not conjugated in general. If we have chosen our origins, then there are four points of special interest, two on each side. These are the infinite points of the object space and the image space, and the origins of coordinates in the object and image space.

In our choice of coordinates these four points are given respectively by $\sigma=0$, $\sigma'=0$, or $h=0$ or $h'=0$.

All magnitudes which relate to the infinite point we shall designate by subscript ∞; all magnitudes relating to the infinite point of the image side, we shall designate by a superscript ∞. All expressions relating to the origin of coordinates in object and image space will have a subscript or superscript 0.

Two of these expressions are well known in optics and a different kind of letter is used for them.

Consider the infinite point in the object space. Its image will be the focal point. If we look at the focal point through our optical system, its apparent distance $\bar{h}/\sigma'=g_\infty'$ is obviously independent of the position of the object origin. This magnitude is a characteristic property of our system. We call it the *focal length* and denote it usually by the letter $f'=g_\infty'$. In the same way we define $\bar{f}=\bar{g}^\infty$ as object *focal length*.

Definition. The *focal length* is the apparent distance of the focal point on the opposite side

FIG. 8. Two rays, one sorting at the object origin. The other converging to the image origin. Figure to demonstrate the formula of Huygens.

seen from an arbitrary point. There is a relation between the two focal lengths which may be found by applying the fundamental law.

We make the following disposition of two rays: the first goes through the focal point in the object space, the second through the focal point in the image space. For these rays

$$\sigma_1' = 0, \quad \sigma_2 = 0. \tag{2, 7}$$

The fundamental theorem then shows that

$$n' h_1' \sigma_2' = -n \bar{h}_2 \sigma_1. \tag{2, 8}$$

We divide by $\sigma_1 \cdot \sigma_2'$ and get

$$n' h_1' / \sigma_1 = -n \bar{h}_2 / \sigma_2 \quad \text{or} \quad n' \bar{f} = -n f'. \tag{2, 9}$$

Theorem. In an optical system which is not afocal, the two focal lengths have the opposite sign and their ratio is equal to the ratio of the refractive indices of objects and image spaces.

The tetrality principle suggests now an investigation of some other special cases. Again our origins may be arbitrary, but we consider the two rays through the origin. This means that

$$\bar{h}_1 = h_2' = 0. \tag{2, 8}$$

The fundamental equation now gives

$$n' h_1' \sigma_2' = -n \bar{h}_2 \sigma_1. \tag{2, 9}$$

Dividing through $\bar{h}_2 h_1'$ we obtain

$$n' \sigma_2' / \bar{h}_2 = -n \sigma_1 / h_1' \tag{2, 10}$$

or

$$n' / g_0' = -n / \bar{g}^0 \tag{2, 11}$$

or verbally:

FIG. 9. Two rays, one leaving the system parallel to the axis, the other coming from the image origin. Origin on object side in the focal point.

Theorem. The apparent distance of the object origin and the apparent distance of the image origin have opposite signs and their ratio is equal to the ratio of the refractive indices.

This is one of the first general theorems in geometrical optics; it was found by Christian Huygens about 1643 and published in his *Dioptrica* in 1702. The form given here is that given by R. Smith (1728).

We get the Huygens form of this theorem if we choose the second ray, so that $\bar{h}_2$ is equal to h_1', and assume that object and image space have the same refractive indices. Then we get

$$\sigma_1 = -\sigma_2'. \tag{2, 12}$$

We consider our eye at the object origin looking at an object of the extension h_1' in the plane through the image origin. Now we interchange eye and object. Eq. (2, 12) shows:

Theorem of Huygens. If we interchange in an optical instrument the position of object and eye, the apparent magnitude of the object will remain unchanged.

Our tetrality principle suggests two other relations between our expressions.

We choose the first ray through the object origin, the second through the infinite point of the image space, which gives

$$\bar{h}_1 = 0, \quad \sigma_2' = 0. \tag{2, 13}$$

The fundamental equation shows that

$$n' h_2' \sigma_1' = n h_2 \sigma_1 \tag{2, 14}$$

or

$$(n' h_2' / n h_2) \sigma_1' / \sigma_1 = 1$$

or

$$(n'/n) \delta^{\infty'} \gamma_0' = 1. \tag{2, 15}$$

FIG. 10. The image of a finite line element orthogonal to the axis.

We shall need this equation in another form. By reference to Eq. (2, 6) we have

$$\delta^{\infty\prime}=\bar{f}/\bar{z}^{\infty}, \tag{2, 16}$$

which gives

$$(n'\bar{f}/n\bar{z}^{\infty})\gamma_0'=1=-\gamma_0'f'/\bar{z}^{\infty}.$$

Eq. (17) gives the connection between the angle magnification at the origin and the distance γ^{∞} of object origin and object focal point.

$$\bar{z}^{\infty}=-f'\gamma_0'. \tag{2, 17}$$

We find the angle magnification equal to one if we take the point at the distance

$$\bar{z}^{\infty}=-f' \tag{2, 18}$$

from the focal point. This point is known in optics as the *nodal point.*

In the same way, if we choose our rays such that the first goes through the infinite object point, and the second through the image origin, then

$$\sigma_1=0, \quad h_2'=0, \tag{2, 19}$$

$$n'h_1'\sigma_2'=nh_1\sigma_2, \tag{2, 20}$$

$$(n'/n)\delta_{\infty}'\gamma^{0\prime}=1. \tag{2, 21}$$

Eq. (2, 6) gives

$$\delta_{\infty}'=z_{\infty}'/f', \tag{2, 22}$$

so that (21) can be written

$$\gamma^{0\prime}=nf'/nz_{\infty}'=-\bar{f}/z_{\infty}'. \tag{2, 23}$$

Eq. (23) gives the angle magnification for the image origin by means of its distance from the image focal point.

The *nodal point* of the image is given by its distance

$$z_{\infty}'=-\bar{f} \tag{2, 24}$$

from the focal point of the image.

(e) The image of a small surface element

In (d) we have taken our origin in object and image spaces arbitrarily, and have disposed our rays. We shall now let one ray be arbitrary and dispose our origins.

First, we assume that our origins might lie in a pair of conjugate points, and our first ray might connect these two points. That means that

$$h_1=h_1'=0. \tag{2, 25}$$

The fundamental law gives for an arbitrary ray

$$n'h'\sigma_1'=nh\sigma_1, \tag{2, 26}$$

in which case we have

$$\sigma_1'/\sigma_1=\gamma_0'=\gamma^{0\prime}=\gamma_1' \tag{2, 27}$$

and

$$(n'/n)\delta'\gamma_1'=1. \tag{2, 28}$$

Eq. (28) shows that the projection coefficient is the same for every ray

$$\delta'=n/n'\gamma_1'=\beta' \tag{2, 29}$$

and is thus dependent only on the angle magnification of the axis point.

All rays through a fixed point in a plane orthogonal to the axis meet in the object space in a fixed point. We have the

Theorem. Within the limits of Gaussian optics, all points of a surface orthogonal to the axis have a fixed image on any surface orthogonal to the axis through the image point of the point of the axis.* This image is similar, and the magnification β' depends only on the angle magnification of the rays emerging from the axis point.

The magnification β' and the angle magnification of the rays through the axis point are connected by Eq. (29). Eq. (29) can easily be deduced from an equation found by R. Smith; it was announced in the form above for the first time by Lagrange, and later rediscovered and generalized by Helmholtz. It is named in text-

* *Note:* In the Gaussian region, surfaces of any finite curvature are indistinguishable from the tangent plane at the axial point.

FIG. 11. The image of a line element orthogonal to the axis at the focal point.

books of optics after one of these three men, usually according to the nationality of the writer. Attributing this theorem to Huygens, as is often done, is incorrect and results from confusion with the theorem given by Eq. (12).

If we wish to know how the magnification β' changes by changing the origin, we compare (29) with (23) and (17). We get

$$\beta' = z_\infty'/f' = \bar{f}/\bar{z}^\infty. \qquad (2, 30)$$

The points where the magnification is equal to one, are called the *cardinal points* of an optical system. Eq. (3) gives

$$\bar{z}^\infty = \bar{f} \quad \text{and} \quad z_\infty' = f', \qquad (2, 31)$$

in verbal form:

Theorem. On either side of an optical system, the distance between the cardinal point and the focal point is equal to the focal length.

We notice further, that the magnitude δ' is a constant for *all* rays, if and only if we have chosen our origins conjugate to each other. In this case, the projection coefficient is identical with the lateral magnification β'.

We have investigated the image in the neighborhood of a finite point. We will now investigate the image in the neighborhood of a focal point.

We choose our coordinate system so that the object origin is in the focal point. Our image origin may be arbitrary. We determine our first ray by

$$\bar{h}_1 = \sigma_1' = 0 \qquad (2, 32)$$

passing through the focal point. We have the fundamental equation

$$n' h_1' \sigma' = -n\bar{h}\sigma_1 \qquad (2, 33)$$

or

$$\bar{h}/\sigma' = -n' h_1'/n\sigma_1. \qquad (2, 34)$$

Hence, if we hold h constant, we get $\sigma' = \text{constant}$.

Theorem. The rays emerging from a point in the focal plane emerge from the system parallel to each other.

We see further that the quantity $g' = h/\sigma'$ is *constant* for all rays. Eqs. (3) and (8) give

$$g' = -n'\bar{f}/n = f'. \qquad (2, 35)$$

This is a new definition of the focal length.

FIG. 12. The image of an infinite line element in afocal systems.

Theorem. All image planes orthogonal to the axis have the same apparent distance from the object focal point. This distance is equal to the object focal length.

We see further that the apparent distance g' has the same value for all rays and all points on the axis. It is constant if and only if the object origin lies at the focal point.

In the same way we find that all bundles or parallel rays at a given inclination to the axis converge to a point in the focal plane whose height h' is given by

$$h'/\sigma = \bar{g} = -nf'/n' = \bar{f}. \qquad (2, 36)$$

Thus, $\bar{g}$, the apparent distance of the image focus, from every object point, is constant for all rays and all object origins. This value is equal to the focal length in the object space.

Finally, we have to consider the special case where the infinite point corresponds to the infinite point in the image space, that is, when the system is telescopic.

We keep the object and image origin arbitrary, and take as the first ray one which enters and leaves our system parallel to the axis. We have

$$\sigma_1 = \sigma_1' = 0. \qquad (2, 37)$$

The fundamental theorem gives

$$n\sigma h_1 = n'\sigma' h_1' \qquad (2, 38)$$

or

$$(n'/n)\gamma'\delta_1' = 1,$$
$$\delta_1' = \delta_\infty' = \delta^{\infty\prime}. \qquad (2, 39)$$

We see that:

Theorem. The angle magnification $\gamma' = \Gamma'$ and therefore on account of (2, 29) the lateral magnification $\beta' = \mathrm{B}'$ of an arbitrary object is constant, if and only if the system is telescopic. A beam of parallel light leaves the system as a beam of parallel light. The angle magnification Γ' depends on the telescopic magnification B', according to the equation

$$n'\mathrm{B}'\Gamma' = n. \qquad (2, 40)$$

The reader recognizes at once that these four problems could be developed one from another with the help of the tetrality principle.

(f) The coordination of the points of the axis

We have investigated the image of a small surface element orthogonal to the axis. Now we will look for the coordination of the points of the axis.

We first let our system of coordinates be arbitrary, one ray being specially chosen, the other general.

The fundamental equation is

$$n(\sigma_1'h'-\sigma'h_1')=n(\sigma_1\bar{h}-\sigma\bar{h}_1). \qquad (2,41)$$

From this equation by dividing by $\bar{h}\bar{h}_1$ we obtain

$$n(1/z-1/z_1)=n'\delta'\delta_1'(1/z'-1/z_1'), \qquad (2,42)$$

by dividing by $\sigma_1 h$ we get

$$ng'(1-\bar{z}_1/\bar{z})=n'\bar{g}_1(z'/z_1'-1), \qquad (2,43)$$

by division with $\sigma\bar{h}_1$

$$ng_1'(\bar{z}/\bar{z}_1-1)=n'\bar{g}(1-z_1'/z') \qquad (2,44)$$

and at last by dividing by $\sigma\sigma_1$

$$n(\bar{z}-{}_1\bar{z})=n'\gamma'\gamma_1'(z'-z_1'). \qquad (2,45)$$

The formulas (42)–(45) become simpler if we take special cases.

The first formula suggests that we make the origins conjugate points. Then we have $\delta'=\delta_1'=\beta_0'$ equal to the lateral magnification at the origin. Then

$$(n'/z')\beta_0'^2-n/z=(n'/z_1')\beta'^2-n/z_1, \qquad (2,46)$$

which means that the expression on the left is independent of z, and is constant for all values of z. We get the value of this constant if we assume, for instance, that $z_1=\infty$. We get $z_1'=f'\beta'$ and by substitution in (2, 46)

$$(n'/z')\beta_0'^2-n/z=n'/f'\beta_0'=-(n/\bar{f})\beta_0'. \qquad (2,47)$$

As a special case, let us suppose that the origins are the *cardinal points*. We then get $\beta_0'=1$, which means that

$$n'/z'-n/z=n'/f'=-n/f. \qquad (2,48)$$

If the origins lie at the *nodal points*, we find $\beta'=n/n'$ or multiplying (47) by $1/n^2$,

$$1/n'z'-1/nz=1/nf'=-1/n'f. \qquad (2,49)$$

In the same way we will now investigate Eqs. (43) and (47). To make g and g' constant we must place our origins on both sides at the focal points. We have then

$$\bar{g}=\bar{g}_1=\bar{f}, \quad g'=g_1'=f', \qquad (2,50)$$

and Eqs. (43) and (46) simplify to

$$\bar{z}z'=\bar{z}_1z_1'. \qquad (2,51)$$

Thus, the product of the distances of the focus from two conjugate points is constant for all rays. We get the value of this constant from (6). Thus, we get *Newton's* equation

$$\bar{z}z'=\bar{f}f'. \qquad (2,52)$$

Finally, formula (45) simplifies only if we have a telescopic system, for in this case and only in this case do we get

$$z'=z_1'=\Gamma' \qquad (2,52)$$

constant.

Eq. (45) becomes

$$n'\Gamma'^2z'-n\bar{z}=n'\Gamma_1'^2z'-n\bar{z}_1. \qquad (2,53)$$

This formula becomes simpler if we place the origins in conjugate points. Then we get $z'=0$ for $z=0$, which means that

$$n'\Gamma'^2z'=nz. \qquad (2,54)$$

(g) The realization of Gaussian optics

We have developed the laws of Gaussian optics from our formula. This section will show how to derive the properties of an actual system specified by its data.

First, for this purpose it is convenient to have another form for the equations:

The values $\bar{h}$ and $n\sigma$ are given; we will calculate h' and $n'\sigma'$. Within the approximation of Gaussian optics h' and $n'\sigma'$ must be linear functions of $\bar{h}$ and $n\sigma$. We get therefore

$$h'=\kappa\bar{h}+\lambda n\sigma, \quad n'\sigma'=\mu\bar{h}+\nu n\sigma. \qquad (2,55)$$

From (55) we have generally for two rays

$$(h_1'n'\sigma_2'-h_2'n\sigma_1')$$
$$=(\kappa\nu-\lambda\mu)(\bar{h}_1n\sigma_2-\bar{h}_2n\sigma_1). \qquad (2,56)$$

The fundamental law has also the same mean-

ing, as if we had postulated

$$\kappa\nu-\lambda\mu=1. \qquad (2, 57)$$

Eq. (57) is the only condition the fundamental law imposes on the coefficients in (55).

In his *Strahlenoptik*, the author called the Eqs. (55) with condition (57) the equations of *Schleiermacher*.

The meaning of the coefficients is found, if we put in (55), $\sigma=0$ and $\bar{h}=0$, respectively, to

$$\kappa=\delta_\infty', \quad \lambda=\bar{g}_0/n, \quad \mu=n'/f', \quad \nu=n'\gamma_0'/n. \quad (2, 58)$$

We consider some special cases. (a) The origins at *conjugate* points. We get $\lambda=\bar{g}_0=0$, $\kappa=\delta_\infty'=\beta_0'$ and from (57)

$$(n'/n)\beta_0'\gamma'=1, \qquad (2, 59)$$

i.e., the equation of *Lagrange*.

The origins in the *cardinal points:*

$$\kappa=\beta_0'=1, \quad \nu=n'\gamma'/n=1, \quad \lambda=0. \quad (2, 60)$$

The origins in the *nodal points:*

$$\kappa=\beta_0'=n/n', \quad \nu=n'\gamma'/n=n/'n, \quad \lambda=0. \quad (2, 61)$$

(b) The origins on both sides in the *focal points:*

$$\kappa=\nu=0, \quad \lambda=\bar{f}/n, \quad \mu=n'/f', \quad (2, 62)$$

Eq. (57) gives

$$\bar{f}/n=-f'/n'. \qquad (2, 63)$$

(c) The system is *afocal*

$$\mu=0, \quad \kappa=\beta', \quad \nu=(n'/n)\Gamma'.$$

The Eq. (57) shows that

$$(n'/n)\beta'\Gamma'=1. \qquad (2, 64)$$

To calculate the coefficients for a given optical system, we first consider the refraction at a single surface. The origin may be at the vertex of the surface. This point must be the *cardinal point*. Eqs. (55) give for this case

$$h'=h, \quad n'\sigma'=\mu h+n\sigma. \qquad (2, 65)$$

The center of the sphere can be regarded as the nodal point. The rays through it ($\sigma=h/r$) remain unrefracted ($\sigma'=h/r$). That gives

$$\mu=(n'-n)/r. \qquad (2, 66)$$

The equations are therefore

$$h'=h, \quad n'\sigma'=((n'-n)/r)h+n\sigma. \quad (2, 67)$$

If we change the origin, shifting it along the axis a distance d, the transformation equations are of the form

$$h'=\bar{h}+(d/n)n\sigma, \quad n'\sigma'=n\sigma. \qquad (2, 68)$$

It is easy to see that these two transformations in the correct sequence give the solution of the problem.

We can give to this solution a simple form if we use the notation of matrices. We regard $\bar{h}$, $n\sigma$ or h', $n'\sigma'$, respectively, as variables. The coefficient scheme of the transformation in (67) may be written by abbreviation

$$\begin{pmatrix} 1 & 0 \\ (n'-n)/r & 1 \end{pmatrix}=R(r). \qquad (2, 69)$$

The coefficient scheme of (68)

$$\begin{pmatrix} 1 & d/n \\ 0 & 1 \end{pmatrix}=D(d). \qquad (2, 70)$$

R depends only on the radius r of the surface, D only on the length d by which the origin is shifted.

It is usual now to write the Eq. (67) in its matrix form

$$\begin{pmatrix} h' \\ n\sigma' \end{pmatrix}= \quad R(r) \quad \begin{pmatrix} \bar{h} \\ n\sigma \end{pmatrix},$$

resp. (68):

$$\begin{pmatrix} h' \\ n'\sigma' \end{pmatrix}=D(d)\begin{pmatrix} \bar{h} \\ n\sigma \end{pmatrix}. \qquad (2, 71)$$

If we have some transformations to make one after another, we name the matrix of the end transformation as the product of the single matrices in the correct order and write the matrices one after another.

Let the origin in the object space be at a distance s from the first surface. Let $d\nu$, $\nu+1$ be the distance from the νth to the $(\nu+1)$th surface. Let s' be the distance of the image origin from the vertex of the last (kth) surface.

Then we get

$$\begin{pmatrix} h' \\ n'\sigma' \end{pmatrix} = D(s')R(r_\kappa)D(d_{\kappa,\,\kappa-1})R(r_{\kappa-1}) \cdots R(r_1)D(s)\begin{pmatrix} \bar{h} \\ n\sigma \end{pmatrix}. \quad (2, 72)$$

In this way we get the image formation within the limits of Gaussian optics in the simplest manner.

(h) Objections to the usual treatment of Gaussian optics

We have shown in this paper how all the laws of Gaussian optics can be developed from one invariant.

We shall show in the next paper that we can deal with more complicated problems in the same way.

We now give once more all the reasons against the usual method of dealing with Gaussian optics.

First, it must be understood that all the laws of Gaussian optics are really nothing else but equations between differential ratios. They are correct for neighboring rays only so long as it is permissible to give equations between differentials instead of between differential ratios.

The laws of *collinear geometry* give by chance the same laws for the rays in the neighborhood of the axis. But this is accidental, and results from the fact that sine and tangent have the same first term in their development. It has been shown that the laws of *collinear image formation* are generally in contrast with optical laws. For instance, if we have a sharp image of a plane, the *sines* of the angles of all rays with the axis must be proportional (Abbē) while collinear geometry would imply the proportionality of the *tangents*.

The idea that we should regard collinear optics as the ideal image formation, and deal with real optics as if the deviation from collinear optics should be regarded as an image error, is erroneous. If it is still said in the optics textbooks that collinear optics forms the skeleton of real optics, it can be said that the authors are mistakenly substituting collinear optics of the *space* for the laws of correct image formation of a *plane*. For, although generally speaking, the whole object space cannot have an exact optical image, yet it is possible that a single surface can be sharply delineated. The laws of this sharp image formation must be regarded as the ideal image formation for this purpose, and the aberrations of a real system must be calculated as a deviation from *this* ideal which is very different from collinear optics.

A New Recording Spectrophotometer

ARTHUR C. HARDY, *Massachusetts Institute of Technology*
(Received June 13, 1935)

A PREVIOUS communication[1] described an instrument designed to measure the spectral reflection or spectral transmission characteristics of materials. The principle underlying the operation of this instrument was substantially as follows: Beams of white light reflected alternately from (or transmitted alternately by) the sample and standard were sent through a prismatic dispersing system, and the light from a narrow spectral region was caused to fall on a photoelectric cell. The intensity of the standard beam was then either increased or decreased by means of a calibrated shutter until the pulsations of the photoelectric current ceased. The shutter was controlled automatically by means of a motor actuated by the amplified photoelectric current. A pen, coupled mechanically to the shutter, recorded the reflection (or transmission) factor of the sample for this spectral region on a sheet of coordinate paper wrapped around a drum. The drum was rotated continuously and was mechanically connected to the mechanism for changing the wave-length. In this way, a curve showing spectral reflection (or spectral transmission) factors throughout the visible range was obtained in from thirty seconds to three minutes, depending upon the type of material being analyzed.

The optical system of this instrument satisfied several requirements that are essential to the attainment of high precision with a photoelectric photometer. The beam from the sample and the beam from the standard fell on the same area of the photoelectric cell, thus eliminating the effect of differences in sensitivity from point to point on the cathode. In addition, both beams illuminated the cathode through the same solid angle, thus eliminating the effect of variations in the response with the angle of illumination. The alternation from one beam to the other was made with such rapidity (60 cycles per second) that changes in the intensity of the source of light, in the sensitivity of the cell, or in the gain in the associated amplifier were of no consequence. Since the instrument was a spectrophotometer, the spectral sensitivity of the cell was unimportant, provided only that sufficient sensitivity was available to operate the device throughout the entire visible region. Satisfying these conditions resulted in a null instrument requiring no calibration of the light source, cell, or amplifier. In contrast with visual spectrophotometers, the reproducibility of a single setting with this instrument depended only upon the amount of amplification that could profitably be used. The speed of operation was vastly superior to that of visual spectrophotometers, and measurements could readily be made at the violet end of the spectrum where visual observations are exceedingly difficult.

Experience with this instrument and with a commercial model[2] based on essentially the same principles suggested several improvements that have resulted in a new instrument with which precision of a higher order is obtained. In the earlier instrument, the transition from one beam to the other was effected by the rotation of a glass disk, alternate sectors of which were silvered and transparent. Regardless of the care with which the disk was prepared, this transition always introduced a large momentary decrease in the photo-cell current at twice the frequency of the signal to be amplified. A similar difficulty was experienced to a lesser degree with the type of flicker device used in the first commercial instrument. In the new instrument, the transition from one beam to the other produces a constant cell current when the two beams are of the same intensity, and produces a pure sinusoidal variation in the cell current when their intensities are unequal. The optical system is such that both beams traverse (except for a split lens between N_2 and the sphere that is not shown in Fig. 2) the same optical parts throughout, thus avoiding any difficulty from selective absorption in the optical parts or from dirt on the various surfaces.

[1] Hardy, J.O.S.A. and R.S.I. **18**, 96 (1929).

[2] Manufactured at the Schenectady plant of the General Electric Company.

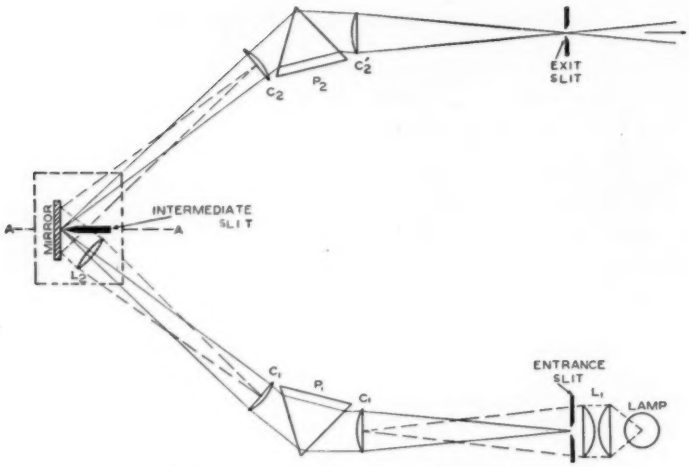


FIG. 1. Optical system of monochromator.

THE MONOCHROMATOR

The dispersing system of the new spectrophotometer is shown in Fig. 1. The source of light is a 6-volt, 18-ampere lamp which is operated from a storage battery.[3] An image of the filament is focused on the collimator lens C_1 by means of the condenser L_1, which is located just in front of the entrance slit of the monochromator. The lenses C_1 and C_1' are so designed that they, in cooperation with the dispersing prism P_1, form a spectrum along the line AA, which is parallel to the axis of the collimator. By means of a first-surface mirror, the light is sent through another similar dispersing system $C_2P_2C_2'$. The lens L_2 images C_1' on C_2. The double monochromator system is used because, for most types of work, freedom from stray light is important.

The wave-length of the light emerging from the exit slit of the monochromator is altered by a motor-driven cam, which moves a carriage supporting the mirror (and also the slit jaw and lens L_2) along the line AA. For simplicity of color specification, the width of the three slits is automatically varied in accordance with the dispersion of the prisms so as to isolate a spectral band having a constant width in millimicrons. This is accomplished by a set of cams, one for each slit, which are operated by the same motor that alters the wave-length. These cams have been so constructed that any desired width of spectral band from one to 10 millimicrons can be selected, and the width of the band is thereafter maintained constant automatically regardless of the wave-length. A dial on the front of the instrument indicates, to the nearest millimicron, the wave-length for which the monochromator is momentarily adjusted. Another dial indicates the width of the slits in millimicrons.

[3] This source has recently been replaced by a monoplane filament projection lamp operated directly from the 110-volt d.c. power supply.

THE PHOTOMETER

The photometering system of the new instrument is represented schematically in Fig. 2. The light leaving the exit slit S of the monochromator passes through a Nicol prism N_1 (or its equivalent) and then through a Wollaston prism W, where it is split into two beams. These beams are

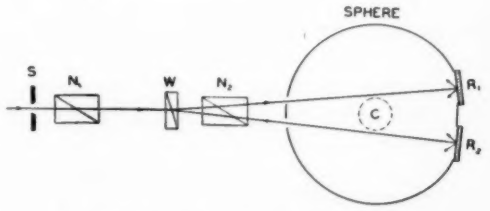


FIG. 2. Optical system of photometer.

polarized in mutually perpendicular planes and pursue slightly divergent courses. Both beams pass through a second Nicol prism N_2, which is mounted inside the hollow shaft of an 1800 r.p.m. synchronous motor. The action of this optical system will be clear from the following analysis. Let the azimuth of prism N_1 with respect to one of the planes of the Wollaston prism be α. Then the amplitudes of the two beams when they emerge from the Wollaston prism are, respectively, $A_1 = A \sin \alpha$ and $A_2 = A \cos \alpha$. If the prism N_2 rotates at a frequency f, the instantaneous values of the amplitudes of the two beams after passing through N_2 are $A_1 = A \sin \alpha \sin \omega t$ and $A_2 = A \cos \alpha \cos \omega t$, where $\omega = 2\pi f$. The corresponding intensities are $I_1 = A^2 \sin^2 \alpha \sin^2 \omega t$ and $I_2 = A^2 \cos^2 \alpha \cos^2 \omega t$. If now the beam of intensity I_1 is reflected into the integrating sphere from a material having a reflection factor R_1, and the beam of intensity I_2 is reflected in a similar manner from a material having a reflection factor R_2, the instantaneous value of the light intensity received by the photoelectric cell located just outside the window C in the sphere wall is proportional to

$$I = I_1 R_1 + I_2 R_2 = A^2 (R_1 \sin^2 \alpha \sin^2 \omega t + R_2 \cos^2 \alpha \cos^2 \omega t).$$

Assuming that $A^2 = 2$, as a matter of convenience, and making a few simple algebraic transformations,

$$I = (R_1 \sin^2 \alpha + R_2 \cos^2 \alpha) - (R_1 \sin^2 \alpha - R_2 \cos^2 \alpha) \cos 2\omega t.$$

Since the photoelectric current is a linear function of the light intensity, at least over a small range, this equation, when multiplied by a suitable proportionality constant, represents the instantaneous value of the cell current. It will be seen that the cell current consists of a steady component proportional to $R_1 \sin^2 \alpha + R_2 \cos^2 \alpha$ and an alternating component of frequency $2f$ and amplitude proportional to $R_1 \sin^2 \alpha - R_2 \cos^2 \alpha$. Since prism N_2 rotates at 1800 r.p.m. the frequency of the alternating component is 60 cycles per second. The alternating component is amplified by means of an amplifier tuned to this frequency, and supplies the field current of a small universal motor whose armature is supplied from another source of the same frequency, actually the 110-volt a.c. power supply.[4] The direction of rotation of the motor depends upon the phase of the amplified photoelectric current, which depends in turn upon whether $R_1 \sin^2 \alpha$ is greater or less than $R_2 \cos^2 \alpha$. By means of a cam and a pair of worm gears, the motor always rotates the Nicol prism N_1 toward the position for which $R_1 \sin^2 \alpha - R_2 \cos^2 \alpha = 0$. At this position, the photoelectric current is constant, the current in the field winding of the motor is zero, and the motor stops. If R_1 is the known reflection factor of a standard material (such as magnesium oxide), the reflection factor R_2 of the unknown specimen is then given directly in terms of α, since the above equation can be written $R_2/R_1 = \tan^2 \alpha$. Hence, this photometer obeys the simple tangent-squared law that is common to many polarization photometers.

A similar analysis applies to the measurement of the transmission factor of transparent materials. If materials with transmission factors T_1 and T_2, respectively, are inserted in the two beams just before they enter the integrating sphere, the result is $T_2/T_1 = \tan^2 \alpha$, provided R_1 and R_2 are equal. In this case there is no need for a material having a known transmission, since T_1 can be made unity by simply omitting the material from the standard beam.[5]

Nicol prisms represent the conventional means of converting natural light into plane-polarized light and, for this reason, their use was assumed in the above analysis. They are not well adapted to the requirements of photoelectric photometry, however, because of the relatively small amount of light flux that they are capable of transmitting. Other things being equal, the amount of flux that can be transmitted through a polarizing prism depends upon both its aperture and its field. Since the field of a Nicol prism is ordinarily small, the aperture must be correspondingly large, and the cost of the prisms becomes prohibitive. Of the various polarizing prisms that were investigated, a modified Rochon prism was found to pass the most light flux for a given amount of calcite. The conventional type of

[4] This is essentially the same mode of operation that was described in the previous communication.

[5] This optical system may also be used to measure the rotation of the plane of polarization, as by a sugar solution. In this case the sample is inserted between N_1 and W.

Rochon prism is composed of two calcite components so arranged that the light travels parallel to the optic axis in one and perpendicular to the optic axis in the other. Although the conventional type of Rochon prism is more effective in its use of calcite than a Nicol prism, it introduces troublesome interference figures. These can be avoided and the amount of calcite can be further reduced by replacing one of the calcite components by glass. A prism of the latter type is illustrated at *B* in Fig. 3, where the first component is glass and the second is calcite cut with the axis perpendicular to the plane of the paper. If the glass has the same index of refraction as calcite for the extraordinary ray, this ray will not be deviated by passage through such a prism. The ordinary ray, on the other hand, will be deviated in the manner shown. By providing sufficient separation between the various elements of the photometer, the ordinary beam can be intercepted by a diaphragm containing an aperture for the passage of the extraordinary beam. This procedure results in a smaller amount of stray light than with Nicol prisms which must absorb the undesired beam on the blackened side of the prism.

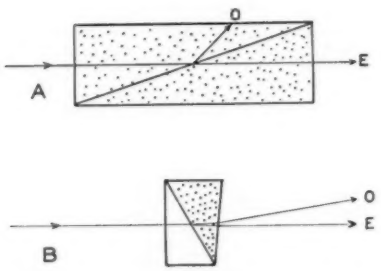


FIG. 3. A Nicol prism *A* of the same optical efficiency as the modified Rochon prism *B* drawn to the same scale.

Since it is impossible to find a glass whose refractive index and also its dispersive power are the same as calcite, the Rochon prisms used in this instrument were achromatized in the following manner: A glass was selected with a dispersive power (ν value = 49.4) the same as that of calcite. The index of this glass for the *D* lines is 1.5696, whereas the index of the calcite for the extraordinary ray is 1.4864. This difference in index was compensated by inclining the exposed surface of the calcite at an angle of 5° 25′ to the exposed surface of the glass. The direction of the extraordinary ray when it emerges from the prism is then substantially parallel to its original direction for all wave-lengths. The slight lateral displacement that the ray suffers is corrected by decentering lenses that are cemented to the glass surfaces. To eliminate the effect of strains in the glass, the Rochon prisms are mounted in the photometer unit with the calcite components facing the Wollaston prism. The saving in the use of calcite is illustrated in Fig. 3, where prism *A* is a Nicol prism (actually a Glan-Thompson prism cemented with Canada balsam) of the same optical efficiency and drawn to the same scale as the modified Rochon prism beneath it.

In the preceding analysis of the behavior of this photometer, it was tacitly assumed that both the sample and the standard had reflection factors that are independent of the state of polarization of the light incident upon them. With this optical system, the incident light is plane polarized, and the plane of polarization rotates at the rate of 30 revolutions per second. This might at first glance be expected to produce the same average result that would obtain if natural light were used. However, the equations previously developed for the instantaneous values of I_1 and I_2 show that they vary with time as $\sin^2 \omega t$ and $\cos^2 \omega t$, respectively. Hence I_1 and I_2 pass through two maxima and two minima (zero) during each revolution of N_2. In general, the white standard has a reflection factor that is independent of the plane of polarization, but test samples are occasionally encountered for which the reflection factor, even for normal incidence, depends to some extent on the plane of polarization. Suppose r is the reflection factor of the sample for the plane in which the intensity of the incident beam is a maximum, and r' is the reflection factor of the sample for light polarized at right angles to this plane. An extension of the above analysis shows that if the values of r and r' are different, a signal of twice the fundamental frequency is generated. The amplitude of this signal is proportional to $r-r'$. This double-frequency signal can be filtered out in the amplifier although, even if it is transmitted by the amplifier, it produces no average torque in the balancing motor provided the armature current of the motor contains no double-fre-

quency component. The analysis shows further that the amplitude of the fundamental frequency is dependent only on r. Hence, the instrument records the reflectance for plane-polarized light with a constant azimuth. If the reflectance for natural or unpolarized light is desired, it can be obtained either by rotating the sample rapidly or by taking the arithmetical average of the values obtained from two readings at azimuths 90° apart. This calculation is valid regardless of the azimuths chosen for measurement.[6]

The Recording Mechanism

The azimuth angle α of prism N_1 can be read to thirty seconds of arc by means of a divided circular scale. However, the rotation of this divided circle is effected by a cam which is driven by the balancing motor, the cam being so fashioned that the rotation of the motor is a linear function of the reflection factor (or transmission factor) of the sample. A dial on the front of the instrument is connected to the gear train ahead of the cam and indicates reflection factors (or transmission factors) directly in percent. Also, a shaft is brought out from this gear train which actuates a pen by means of a lead screw, the total travel of the pen between 0 and 100 percent being 10 inches. Special coordinate paper is used, the horizontal rulings being spaced 0.5 percent apart. This paper has an overall size of $8\frac{1}{2}\times 11$ inches and is accurately registered before each test on a platen which moves linearly with the wave-length. The wave-length scale from 400 to 700 millimicrons occupies $7\frac{1}{2}$ inches and the vertical rulings are spaced 2 millimicrons apart.

Special cams have been constructed to expedite the solution of special problems. One such cam makes it possible to plot optical density directly as a function of wave-length, and special coordinate paper has been printed for use with this cam. Another cam plots log log R or log log T as a function of wave-length, a procedure that is useful in studying materials of unknown thickness or of unknown concentration. If a material obeys Bouguer's law or Beer's law, the shape of the curve obtained with this cam is independent of the thickness or concentration, and the curve is merely shifted vertically on the sheet with variations in thickness or concentration. This cam is especially useful in testing colored glasses of different thicknesses or in identifying solutions of unknown concentrations. Still another cam records reflection factors (or transmission factors) on a linear scale multiplied by 5. This is useful for samples with low reflection or transmission factors. These special cams are so constructed as to be readily interchangeable with the linear cam that is ordinarily used.

Due to the nature of the optical system, this instrument evaluates the amount of light that is diffusely reflected by an opaque material when illuminated by a nearly unidirectional beam falling normally on the surface. Because of the Helmholtz reciprocity law, this is equivalent to illuminating the sample diffusely and measuring the normal brightness relative to that of magnesium oxide. Measurements of transmission are made by using two equal white reflection standards and inserting the material whose transmission is to be measured in one of the beams just before it enters the integrating sphere. The values obtained indicate the diffuse transmission of the sample, although in the case of homogeneous materials, the specular and diffuse transmission factors are substantially equal.

Precision

The cam controlling the motion of the carriage which changes the wave-length has been calibrated at thirty-six points by means of lines in the mercury arc spectrum. If precautions are taken to avoid backlash by operating the platen always in one direction, the probable error in wave-length is less than 0.5 millimicron. When the wave-length is held constant, the balancing motor will reproduce a setting to approximately ± 0.05 percent when the reflection factor or transmission factor is 100 percent. By employing a method of coupling the photoelectric cell to the grid of the first tube that was described on page 105 of the previous communication, the reproducibility improves as the reflection factor decreases. This reproducibility is somewhat better than the precision with which the extinction point of N_1 and W can be determined. In general, however, the error introduced by uncertainty in the setting of the extinction point may be made

[6] R. D. Nutting, J.O.S.A. **25**, 211 (1935).

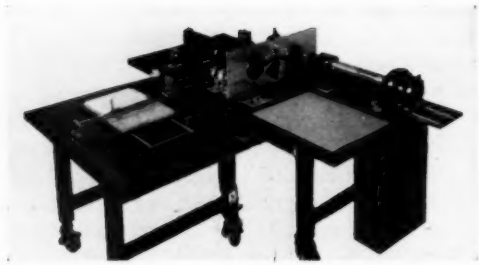


FIG. 4. Photograph of spectrophotometer showing the recording table at the left, the monochromator in the center background, and the photometer and integrating sphere at the extreme right. Reading from left to right, the dials indicate the wave-length in millimicrons, the slit width in millimicrons, and the reflection factor (or transmission factor) in percent.

less than 0.10 percent. Values of reflection factors are dependent, of course, upon the white standard that is used. It has been found that the most satisfactory white standard is a block of magnesium carbonate which has been smoked with magnesium oxide. Two such standards are always prepared, and they are compared, one against the other in normal and reversed positions, to insure that they are alike at all wave-lengths to at least 0.5 percent.

When the instrument is in operation, there is a possibility that the balancing motor will be unable to follow rapid changes in reflection and transmission factors because of the inertia of the various parts, principally the armature of the motor. For this reason, the motor that alters the wave-length has been provided with a speed control. The operator soon learns to adjust the speed of the wave-length motor in such a manner as to enable the balancing motor to maintain prism N_1 sufficiently close to the true point of balance. In general, a satisfactory curve for a nonselective reflector can be drawn in less than one minute, whereas a satisfactory transmission curve of a didymium filter that possessed many sharp absorption bands might require eight to ten minutes. Eventually, it is expected that an automatic speed control will be installed. A photograph of the instrument is shown in Fig. 4.

THE INTEGRATING ATTACHMENT

An integrating attachment was described in connection with the earlier instrument with which it was intended to perform the three integrations that are necessary in obtaining the tristimulus values of a material for a standard light source and observer. The construction of this attachment was postponed pending the completion of the new spectrophotometer. In the meantime, a simpler method of mechanical integration has been devised, and an attachment which integrates by the selected-ordinate principle has been constructed.[7] A photograph of the integrating mechanism with the cover removed is shown in Fig. 5.

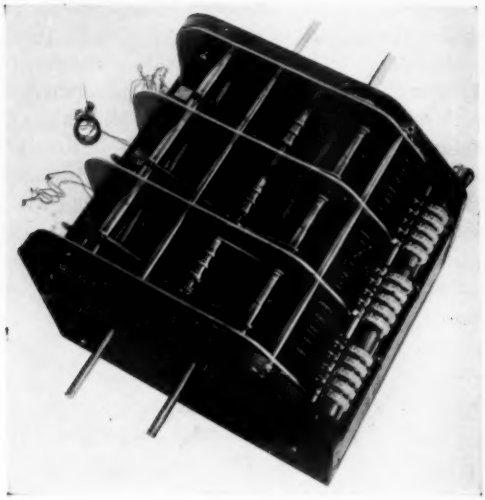


FIG. 5. A photograph of the integrating attachment with the cover removed.

In applying the selected-ordinate principle, the I.C.I. distribution functions for a standard illuminant and standard observer are integrated once and for all, and each of the three integrals is divided into 100 parts. The wave-lengths which would divide each integral into 100 equal parts are then ascertained. Contacts are made as the spectrophotometer arrives at each of these wave-lengths by means of a commutator in the form of a drum containing three sets of pins located along a spiral path. The integrating attachment itself consists of three separate adding machine

[7] As this paper is submitted for publication, the integrator has been constructed but has not been attached to the spectrophotometer because of a few changes that must be made in the mechanical system of the spectrophotometer.

units constructed largely from standard Marchant parts. The key mechanism that is customarily operated manually is replaced in this case by a set of step cams which are controlled by the shaft that drives the recording pen. Whenever one of the three brushes makes contact with a pin on the commutator drum, the resulting current actuates a solenoid which sets into operation one of the adding machine units. This operation adds into the proper totalizing carriage the reflection factor (or transmission factor) of the sample at this wave-length. At the end of the run, the three dial readings, when multiplied by appropriate factors, indicate the tristimulus values of the material. These readings can be readily converted into dominant wave-length, brightness, and purity by means of charts which are now in preparation. The construction of the integrating attachment was made possible by a grant from the Rumford Fund of the American Academy of Arts and Sciences.

Pending the completion of the integrating attachment, tristimulus values have been computed manually by the selected-ordinate method, using either 10 or 30 ordinates as the nature of the problem required. For the I.C.I. Illuminant C and the standard I.C.I. observer, the appropriate wave-lengths are given in the accompanying table. The operation consists in adding the reflection factors (or transmission factors) at the wave-lengths indicated in each of the three columns. Each sum is divided by the number of ordinates used and multiplied by an appropriate factor, which takes account of the fact that I.C.I. Illuminant C does not fall at the center of the I.C.I. triangle. These multiplying factors are listed in the table at the foot of each column. The obtaining of these summations is greatly simplified by preparing three glass plates, each of which is ruled with vertical lines corresponding to the wave-length scale of the standard coordinate paper used with the spectrophotometer. Using these plates and an adding machine to perform the summations, an experienced operator can obtain the three necessary summations, using 30 ordinates, in approximately six minutes. (See Table I.)

TABLE I. *List of 30 selected wave-lengths (millimicrons) for I.C.I. Illuminant C and Standard Observer.* (The ten starred ordinates yield results of sufficient precision for materials that are not highly selective.)

X	*Y*	*Z*
424.4	465.9	414.1
*435.9	*489.4	*422.2
443.9	500.4	426.3
452.1	508.7	429.4
*461.2	*515.1	*432.0
473.9	520.6	434.3
531.0	525.4	436.5
*544.2	*529.8	*438.6
552.3	533.9	440.6
558.7	537.7	442.5
*564.0	*541.4	*444.4
568.8	544.9	446.3
573.2	548.3	448.2
*577.3	*551.7	*450.2
581.2	555.1	452.1
585.0	558.5	454.0
*588.7	*561.9	*455.9
592.3	565.3	457.9
595.9	568.8	459.9
*599.5	*572.5	*462.0
603.2	576.4	464.1
606.9	580.4	466.4
*610.8	*584.8	*468.8
614.9	589.5	471.4
619.2	594.8	474.4
*624.0	*600.7	*477.8
629.6	607.6	481.9
636.4	616.0	487.3
*646.2	*627.1	*495.3
662.2	647.0	511.5
Multiplying factor:		
0.9804	1.000	1.1812

The author desires to express his appreciation to Mr. O. W. Pineo and to Dr. D. G. C. Luck for their assistance in designing and assembling this instrument, and to Mr. Carl G. Selig who constructed most of the mechanical parts. The author is also indebted to Dr. R. D. Nutting, Mr. D. L. MacAdam and Mr. R. H. Zinszer for many suggestions for improvements during the two years or more that the instrument has been in operation. The construction of the integrator is due almost entirely to the efforts of Mr. E. A. Nordberg, whose assistance is gratefully acknowledged.

Some Stigmatic Grating Mountings

HAROLD T. SMYTH, *George Eastman Research Laboratory of Physics, Massachusetts Institute of Technology*

(Received June 3, 1935)

The general conditions for the production of stigmatic images with concave or plane gratings are derived. Several practical methods of realizing these conditions have been devised using a cylindrical lens with plane mirrors, or a concave mirror alone, between the slit and the grating.

General Theory

FOR intensity measurements or interferometric work it is sometimes desirable to combine the stigmatic properties of a prism instrument with the high dispersion and resolving power of a large grating. Concave grating mountings are in general not stigmatic but can be made so over a limited range by the use of cylindrical lenses or concave mirrors, as has been done in the Wadsworth mounting and the various mountings suggested by Runge and Mannkopff[1] and by Oldenberg.[2]

Allowing for the possibility of having such a system of lenses or mirrors between the slit and the grating, we shall, in order to be quite general, use different symbols for the distances of the horizontal and vertical images of the slit from the grating, calling these, respectively, r_1 and r_2. The light strikes the grating at angle θ with the normal. Let the horizontal focus of a spectral line corresponding to a reflection angle of θ' be at distance r_1' from the grating and the vertical focus at distance r_2'. Then,[3]

$$\left(\frac{\cos^2\theta}{r_1}-\frac{\cos\theta}{\rho}\right)+\left(\frac{\cos^2\theta'}{r_1'}-\frac{\cos\theta'}{\rho}\right)=0 \quad (1)$$

and

$$\left(\frac{1}{r_2}-\frac{\cos\theta}{\rho}\right)+\left(\frac{1}{r_2'}-\frac{\cos\theta'}{\rho}\right)=0. \quad (2)$$

For fixed values of r_1, r_2 and θ the loci of r_1' and r_2' are curves symmetrical about the normal to the grating. To make these curves touch without intersecting we must make $r_1'=r_2'$ when $\theta'=0$.

This when substituted in (1) and (2) gives immediately that

$$\cos^2\theta/r_1=1/r_2 \quad \text{or} \quad r_1/r_2=\cos^2\theta, \quad (3)$$

a condition independent of ρ, the radius of curvature of the grating.

If (3) is satisfied then again from (1) and (2) we get that

$$\cos^2\theta'/r_1'=1/r_2' \quad \text{or} \quad r_1'/r_2'=\cos^2\theta',$$

which shows how the astigmatism increases as we observe the spectrum away from the normal.

Use of Cylindrical Lenses and Plane Mirrors

To fulfill the conditions for a stigmatic image on the normal one can use a Rowland or Abney mounting in which is placed a cylindrical lens between the slit and the grating, in such a position that a virtual vertical image of the slit is formed on the tangent drawn to the Rowland circle at the center of curvature of the grating. It can very easily be shown that condition (3) is satisfied by such an arrangement.

By means of a single plane mirror a Paschen mounting can be modified to give different regions on the normal and at the same time keep the slit at a fixed position. In Fig. 1, S is the slit, G the grating, M the mirror, C the center of curvature of the Rowland circle and N the center of curvature of the grating. Let GM produced meet the Rowland circle at P. The

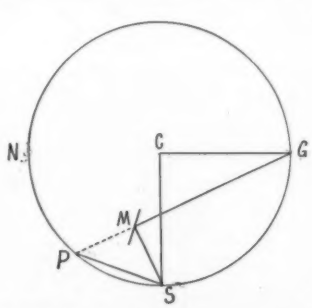


FIG. 1. Single plane mirror mounting.

[1] Runge and Mannkopff, Zeits. f. Physik **45**, 13 (1927).
[2] Oldenberg, J.O.S.A. **22**, 441 (1932).
[3] Kayser, *Handbuch der Spectroscopie*, Vol. I, p. 456.

position of the mirror must be such that $SM = PM$. Hence $\angle MPS = \angle MSP$ and $\angle GMS$ = twice $\angle GPS$. But the fixed $\angle SCG$ = twice $\angle GPS$. Therefore $\angle GMS = \angle SCG$, so that M moves on the circle through SCG. It is obvious that the plane of the mirror goes through the fixed point C. A cylindrical lens between S and M can be used to give a stigmatic image at N since the mounting is now effectively a Rowland mounting.

Since this mounting involves the necessity of having the light come through the slit S in different directions for different settings it is perhaps preferable to use two mirrors. The following arrangement of the two mirrors was found to be very satisfactory.

The two mirrors A and B (Fig. 2) were mounted on a carriage C which slides along the tracks T. Both mirrors can be rotated about a vertical axis in the plane of the reflecting surface and while B is fixed in position on the carriage A can be slid backwards and forwards on the carriage any required amount by means of the screw L. The position of the carriage on the tracks determines the position of B which in turn determines the angle of incidence on the grating and hence the wave-length falling at N. Focusing, which is done empirically by moving A, is accomplished when $SA+AB+BG=PG$. A cylindrical lens between S and A is used to produce a stigmatic image at N.

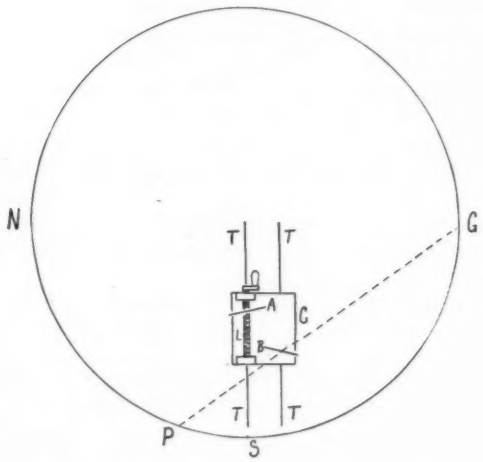


FIG. 2. Double plane mirror mounting. The size of the carriage has been greatly exaggerated to bring out the details.

USE OF CONCAVE MIRRORS

Runge and Mannkopff[1] have given a neat geometrical construction to determine the position of the horizontal and vertical foci of a concave spherical mirror when illuminated from a point off the normal. We are interested only in cases where the source is on the Rowland circle for the mirror. The horizontal focus will therefore lie also on the Rowland circle, and the vertical focus, which lies on the line joining the source to the center of curvature of the mirror, will be real or virtual depending on whether the angle of incidence is greater or less than 45° as illustrated in Fig. 3a and b. S is the source, M the mirror, C the center of curvature, H the horizontal focus and V the position of the vertical focus which is a horizontal line lying along SCV.

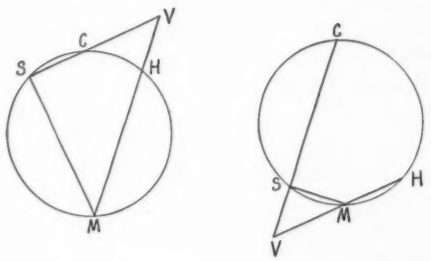


FIG. 3. Diagram showing the position of the vertical and horizontal foci of a concave mirror where the angle between the incident ray and the normal is (a) less than 45° and (b) greater than 45°.

The arrangement in Fig. 3b can be used to give stigmatic images of spectral lines on the normal of a concave grating. In Fig. 4 GPN is the Rowland circle of a concave grating G to which GN is normal. NM bisects the angle PNS and meets GP in M. MS is perpendicular to NS the tangent to the Rowland circle at N. At M is placed a mirror whose center of curvature is at N. Light coming through a slit at S in the direction SM will form a horizontal image of the slit at P and a virtual vertical image at V. The result will be that the spectrum at N will be stigmatic.

The conditions (3) are independent of the

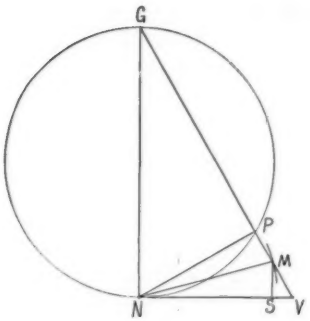


FIG. 4. Arrangement of a concave mirror to produce stigmatic images on the normal of a concave grating.

radius of curvature and can be applied to a plane grating where r_1 and r_2 are, respectively, the distances behind the grating of the horizontal and vertical images formed by a concave mirror arranged as in Fig. 3a. The following arrangement suggests itself. In Fig. 5, S is the slit, M a concave mirror with center at C where the angle MSC is a right angle. The line CGP which is normal to the grating G and the photographic plate P is parallel to SM. The grating lies at the intersection of CP and the perpendicular bisector of CM. It is easily shown that the line CM bisects the angle SMG so that if the grating were not in the way the horizontal and vertical images of S would fall along MG at V and H, respectively, where CH is perpendicular to MV.

If the angle $MGP=\theta$, then $HG=CG\cos\theta=VG\cos^2\theta$ satisfying the conditions for stigmatic images on the grating normal. Putting $\rho=$infinity in either (1) or (2) gives that the vertical and horizontal images formed by the grating will lie together at P where $GP=GV$, making $\angle GMP$ a right angle. The locus of the

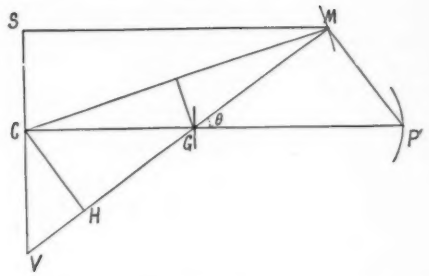


FIG. 5. Plane grating mounting with single concave mirror.

horizontal focus, as pointed out by Monk,[4] is not a circle but a lemniscate. The radius of curvature of this curve at P is $1/3\ GP$. The locus of the vertical focus is a circle through P with G as center.

This mounting has been set up for one region with a concave mirror of 17 ft. radius of curvature and a plane grating (15,000 lines per inch) and has sufficient resolving power for intensity work with a sector disk or a stepweakener at the slit. For such a purpose a region of at least 1500A in the first order can be covered in one exposure. It would be possible by not too complicated mechanical means to make the mounting completely or at least partially automatic in focusing so that change of region could be accomplished with a minimum of trouble.

ACKNOWLEDGMENT

The author wishes to acknowledge the valuable suggestions and helpful criticism of Professor G. R. Harrison under whose direction this work was done.

[4] Monk, J.O.S.A. **17**, 358 (1928).

A Modification of the Fabry-Perot Interferometer

R. C. MACHLER AND RUSSELL A. FISHER, *Northwestern University*

(Received June 3, 1935)

A Fabry-Perot interferometer is described in which one of the plates is silvered to opacity and the film then ruled with parallel slits. It is demonstrated that by this means either the resolving power or the intensity of the fringes can be materially increased as compared with the conventional interferometer. The influence of the rulings upon the resolving power and intensity is discussed and the possible spurious effects considered. A manner in which the film may be ruled without damage to the plate is described. The question of the appropriate width and spacing of the rulings is also discussed.

INTRODUCTION

THE resolving power of a Fabry-Perot interferometer is proportional to the order of interference and to a function of the reflection coefficients of the metallic films. The ratio of the resolving power to the order, which may be termed the resolving power per order,[1] expresses the number of equal parts into which the interval in retardation between successive orders may be resolved. It was shown by Childs[2] that for an interferometer with perfect optical surfaces

$$\frac{\text{Resolving power}}{\text{Order}} = \frac{\pi R^{\frac{1}{2}}}{1.21(1-R)}. \qquad (1)$$

Here R is the geometric mean of the reflection coefficients of the two semi-transparent metallic films.

A large resolving power per order is desirable in the analysis of fine line structures for in this way high resolution is attainable with a minimum of order overlapping. Since silver forms the most highly reflecting films for visible light our discussion will refer entirely to silver films. Although the resolving power of the interferometer increases rapidly with the reflection coefficients of the films the maximum reflecting power of which silver is capable can never be employed because the fraction of incident light transmitted by the metallic film becomes vanishingly small before the maximum reflecting power is attained.[3]

We have found that a distinct gain in resolving power per order can be accomplished without sacrifice in intensity through the use of an interferometer one film of which is silvered to opacity, thus attaining the maximum reflection coefficient of nearly 94 percent for green light. Light is introduced through the opaque film by means of narrow parallel slits produced by the removal of strips of silver from the film. The second film is left slightly transparent. It will be shown that the spacing of the slits in the opaque film may be so chosen that an interferometer thus modified has approximately the same resolving power as would a conventional instrument with films of the same mean reflection coefficient.

Light which enters a slit in the first film at an appreciable angle with the axis suffers multiple reflections as is indicated in Fig. 1, the light

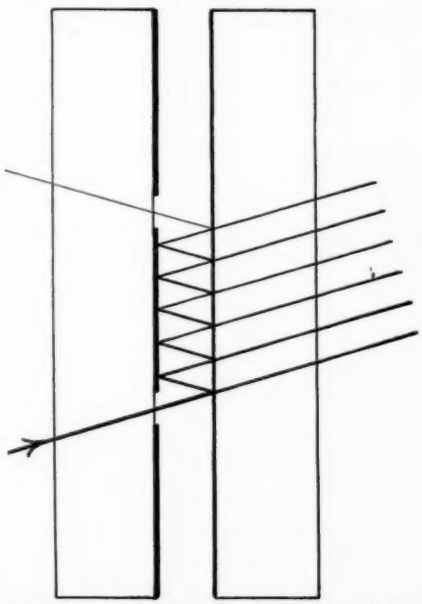


FIG. 1. Multiple reflection of light between silver films; first film opaque and slitted, second film semi-transparent.

[1] This quantity is sometimes termed the "effective number of reflections."

[2] Childs, J. Sci. Inst. **3**, 97 (1926).

[3] See, for example, experimental data of Murmann, Zeits. f. Phvsik **80**, 161 (1933).

emerging from the second film being in a state of interference. The train of reflections is eventually terminated by escape of the remaining light through a neighboring slit, but if the slits are appropriately spaced, the light so lost is negligible. This interference has some features in common with that occurring in a Lummer plate. It also resembles that in a modification of the Fabry-Perot interferometer proposed by Barnes[4] in which the light is introduced at the edge of the plates.

Fig. 2 is an enlarged photograph (negative) of the interference pattern for the mercury line λ5461 given by the ruled interferometer with a plate separation of 2.35 mm. The distance between rulings was 4 mm and the width of the rulings 0.1 mm. The resolving power per order attained in this pattern is more than 30. This requires a mean reflecting power of more than 92 percent according to Eq. (1). It is to be noted that this pattern does not have circular symmetry but that the fringes are greatly weakened along a line through the center of the ring system which is parallel to the rulings.

Resolving Power

The intensity I at any point in the fringe system formed by a conventional Fabry-Perot interferometer is given by the well-known equation

$$I = I_0 T_1 T_2 1/((1-R^2)+4R\sin^2(\beta/2)), \quad (2)$$

where I_0 is the intensity in the absence of the interferometer, T_1 and T_2 are the transmission coefficients of the first and second films, respectively, and β is the phase retardation between successive interfering contributions. This intensity equation is obtained through a summation of the infinite series of amplitude contributions arising from the multiple reflections, the phase differences being taken into account. The resolving power expression, Eq. (1), is obtained from Eq. (2) by determination of the minimum difference in β permitting the resolution of two equal fringe maxima in accordance with Rayleigh's criterion; i.e., that two maxima are resolved when there exists between them a minimum of height less than 0.81 times that of one of the maxima alone.

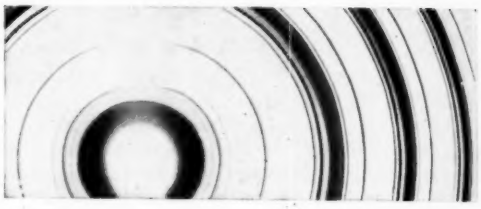


Fig. 2. Interference pattern of mercury line λ5461 obtained using ruled interferometer having plate separation of 2.35 mm.

We obtain a corresponding intensity equation for the ruled interferometer by summation of the finite number of amplitude contributions arising from N double reflections which occur in the terminated series. The intensity equation is

$$I = I_0 F T_2 \frac{(1-R^N)^2+4R^N\sin^2(N\beta/2)}{(1-R)^2+4R\sin^2(\beta/2)}. \quad (3)$$

Here F is the fraction of incident light passing through the slits in the first film.

If we plot the I *versus* β relation for the two types of interferometer, both having the same value of R, on the same graph, we obtain such curves as are given in Fig. 3. The two curves are reduced to the same scale by representing the intensity in percent of its maximum value. The solid curve is for a ruled interferometer with $N=16$ and the dashed curve for a conventional interferometer. R has the value of 0.90 in both curves. It may be seen from the equations that the curve for the ruled interferometer touches that for the conventional interferometer at points for which $\sin N\beta/2=0$, and rises to form small secondary maxima at points for which $\sin N\beta/2=1$. There are N such secondary maxima between two primary maxima and their height is $1+4R^N/(1-R^N)^2$ times that of the conventional curve.

An explicit expression for the resolving power of the ruled interferometer obtained from the intensity Eq. (3) contains a slowly converging series of terms and is not given here since it is not of great use. It may be observed from a comparison of the intensity Eqs. (2) and (3) and a comparison of the curves of Fig. 3 that the resolving power of the ruled instrument ap-

[4] Barnes, Astrophys. J. 19, 190 (1904).

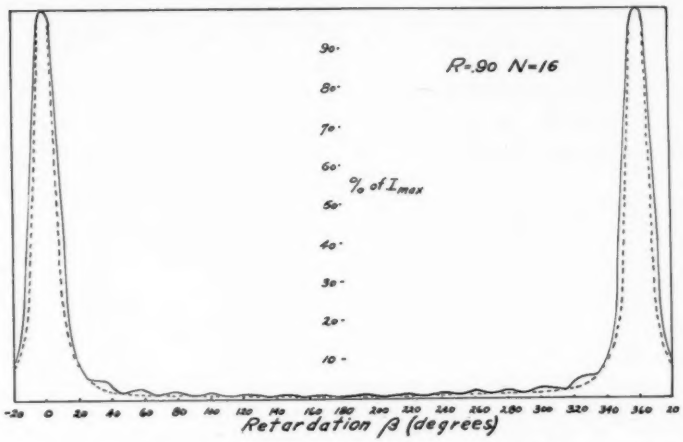


FIG. 3. Comparison of intensity *versus* retardation curves for conventional and ruled Fabry-Perot interferometers. The dashed curve refers to the conventional instrument, the solid curve to a ruled instrument allowing 16 reflections. The mean reflection coefficient is 0.90 in both cases.

proaches closely that of the conventional instrument as R^N becomes small in comparison with unity.

It is hardly to be expected that the theoretical resolving powers can be attained in practice with either instrument since in the derivations the assumption of perfect optical surfaces is made.[5] However, irregularities in the surfaces would introduce greater discrepancies when the long trains of reflections of the conventional interferometer are involved than when the trains are terminated as in the ruled instrument.

We have compared the resolving powers of the ruled and unruled interferometers directly by ruling one-half of a dense but not opaque film and leaving the other half unruled. Photographs of the interference pattern of the mercury line λ5461 using first the ruled and then the unruled portions of the films were made under the same conditions. In all cases the resolving power of the ruled instrument equaled that of the unruled instrument if the slits were so spaced as to make N as great as 15. When dense films were used requiring long exposures with the unruled interferometer the ruled instrument always surpassed it in resolving power. This is probably to be attributed to instabilities in the apparatus which operate to reduce the effective resolving power in long exposures.

The secondary maxima represented in Fig. 3 have been observed only when the number of reflections was as small as 5 or less. This is to be expected since when N is larger the secondary maxima become extremely faint. In addition N is probably not the same for all contributions to a given point in the fringe system.

INTENSITY

It is to be seen from (2) and (3) that the intensity in fringe maxima for the conventional interferometer is given by

$$I_{\max}=I_0T_1T_21/(1-R)^2, \qquad (4)$$

and for the ruled interferometer by

$$I_{\max}=I_0(T_2/N)(1-R^N)^2/(1-R)^2. \qquad (5)$$

Here F from (3) is replaced by $1/N$ since it is evident from the geometry (Fig. 1) that for a point in the fringe system for which the number of reflections occurring before escape of light through an adjacent slit is N the fraction of incident light admitted is approximately $1/N$.

The gain of intensity of the ruled instrument over the conventional instrument is represented by the substitution of the quantity $(1-R^N)^2/N$ for T_1, the transmission coefficient of the first

[5] Burger and Van Cittert, Zeits. f. Physik **44**, 58 (1927), have made a careful investigation of this departure from the theoretical resolving power.

film, which is vanishingly small for films reflecting more than 90 percent of the incident light. The fraction $(1-R^N)^2/N$ varies slowly with N when R is of the order of 0.9 so that N may be made large enough to permit high resolving power without materially reducing the value of the fraction.

If high resolving power per order is not desired so that both films may be left semi-transparent there is still an increase in intensity of several fold to be gained by ruling the first film.

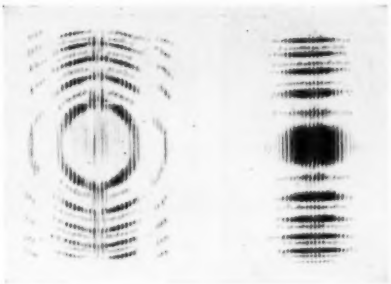


FIG. 4. Diffraction effects, obtained with ruled interferometer using collimated beam from slit source, (a) light first incident on ruled film, (b) light first incident upon unruled film.

DIFFRACTION EFFECTS

Two distinct diffraction effects are to be expected: first, there is the spreading of the light at passage through a narrow slit in the opaque film and second, the interference of the contributions from the several equally spaced slits after the manner of a diffraction grating. Both of these effects can be demonstrated experimentally, and both can be shown to have no effect upon the circular interference fringes when the light is incident from the slitted side of the interferometer.

Fig. 4 gives photographs illustrating the possible diffraction effects. Fig. 4a shows the pattern formed when the interferometer is placed in the collimated beam from a narrow slit, the light entering the interferometer from the side of the ruled film, and the rulings being parallel to the slit. The light is seen distributed into the familiar Fraunhofer diffraction pattern with broad central maximum and successive secondary maxima formed by passage of the light through the narrow rulings of the film. Superposed upon this pattern are numerous fine maxima produced by the rulings acting as a coarse diffraction grating. The circular interferometer fringes are seen crossing the diffraction pattern but undistorted by the diffraction effects. The distance between rulings was in this case 1 mm and their width 0.1 mm. Fig. 4b shows the pattern formed in the same optical system with the interferometer reversed so that the light is first incident upon the unruled film.

When the light is first incident upon the ruled film the diffraction phenomena cannot affect the interference fringes since the multiple reflection takes place subsequent to the spreading of the light. The only effect of diffraction therefore can be a spreading and redistribution of the light in the rings. If, however, the light enters the interferometer through the unruled film and emerges from the narrow slits after multiple reflection there is a spreading such that all light traveling in a given direction is no longer in the same state of interference. Thus the interference rings are destroyed.

It is obvious that if an extended source is substituted for the slit source used in the arrangement producing 4a the diffraction minima will disappear and the complete ring system will be formed.

PREPARATION OF THE RULED FILM

The removal of fine strips of silver film without injury to the surface of the interferometer plate requires extreme care. The silver films which we have used are laid down by evaporation in high vacuum. Such films are so soft during the first few hours after being deposited that they may be removed by the lightest touch. We have found a most satisfactory means for removing the film to be a light stylus of ivory mounted on the ruling arm of a dividing engine and so balanced as to rest very lightly upon the surface to be ruled. Ivory is easily carved to required fineness and insures careful handling because of its own fragility. Such a stylus when carefully handled will rule uniform lines of less than 0.1 mm width. We are unable to detect damage to a surface ruled numerous times by such an ivory stylus.

The most suitable width and spacing for the rulings depends upon the mean reflection coefficient of the films and upon the distance between the plates. Our experience indicates that it is a satisfactory rule to make the number of double reflections as large as the resolving power per order to be expected from the films according to Eq. (1). Thus N will usually be between 15 and 40. It can readily be shown from the geometry that the distance S between rulings which will give N double reflections for contributions to the rth ring of the pattern counting outward from the center is,

$$S=2N(\lambda r d)^{\frac{1}{2}}. \qquad (5)$$

Here d is the distance between the plates and λ is the wave-length of the light. For most purposes it is well to refer to the second ring making r equal to 2. Since the width of the rulings should be $1/N$ times the spacing the appropriate widths are also indicated by (5).

When plate separations of several centimeters are used S may be more than a centimeter and under these circumstances only one relatively wide ruling may be desirable. In other cases where intensity is not a consideration the terminated series of reflections may be avoided by the use of only one ruling. Since the intensity in the fringes is proportional to the plate area employed it is usually desirable to rule the entire plate surface.

In order to facilitate adjustment of the instrument when the slit spacing is large it is desirable to make one ruling transverse to the several parallel ones and near the edge of the plate. This ruling may be excluded by a stop when the instrument is in use.